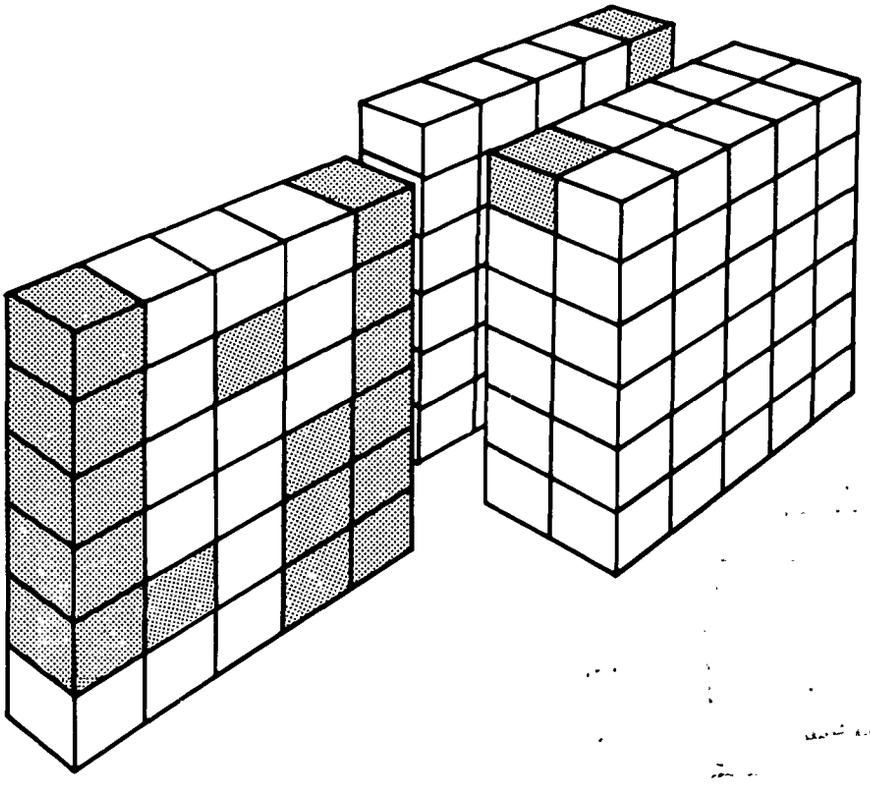
A FACTOR ANALYSIS OF THE SYMBOLIC-EVALUATION ABILITIES

Studies of Aptitudes of High-level Personnel

R. Hoepfner

J. P. Guilford

P. R. Merrifield



REPORTS FROM THE PSYCHOLOGICAL LABORATORY

THE UNIVERSITY OF SOUTHERN CALIFORNIA

Number 33 December 1964

TABLE OF CONTENTS

LIST OF TABLES

A FACTOR ANALYSIS OF THE SYMBOLIC-EVALUATION ABILITIES[1]

THE PROBLEM

Attempting to understand human intellectual processes, philosophers and psychologists have traditionally set the goal of achieving a conceptual system embracing all intellectual activities. Within the lists of intellectual concepts are such words as decision, estimation, evaluation, hypothesizing, judgment, and verification. All of these thought processes have one thing in common; they produce no new information; they only operate upon cognized or produced information and result in a decision. They serve the function of stabilizing an uncertain state of affairs or of clarifying a somewhat unstructured situation.

The stabilizing and clarifying functions of these judgmental processes made them amenable to early experimental investigation: subjects merely had to choose between two or more specified courses of action when presented with a stimulus to be judged. The results of judgment experiments were similar to those of the early reaction-time and perceptual experiments in one respect; the main effects of the experimental variables were often obscured by large within-group response variability. Whereas individual differences very early became a problem in connection with reaction-time experiments, and resulted in broadened investigation in the perceptual area, in the form of psychophysics, it had little effect upon later attempts to study judgmental processes empirically.

The present study is an attempt to approach empirically the problem of describing judgmental processes from the standpoint of individual differences, as a step toward understanding those processes. It attempts to identify basic traits with respect to which individuals differ in judgmental performances. Thus, for this study, individual differences in judgmental tasks are considered in terms of traits, more specifically, intellectual abilities.

In 1960, Guilford and Merrifield described a theoretical model called the structure of intellect, based upon a number of intellectual abilities identified in past factor-analytic studies. The model offers a definition of intelligence that is free from the operational, but circular, definition of intelligence offered by Boring (1923) that "Intelligence is what the tests test." In the structure-of-intellect model, the primary intellectual abilities are classified in terms of three dimensions: operations, contents, and products. There are five operations: cognition, memory, divergent production, convergent production, and evaluation, which the operator can perform upon information. There are four kinds of contents, or categories of information, upon which the organism can perform the operations. The four contents are called figural, symbolic, semantic, and behavioral. The third dimension of the model represents the six kinds of products: units, classes, relations, systems, transformations, and implications. The products are the results of the operator's psychological processing of information. The intersection of a given operation, content, and product implies a unique factor of intelligence. This results in the model's prediction of 120 abilities, 61 of which had received some empirical verification before the results of this study became known.

This study selected for investigation the six factors of symbolic evaluation predicted from the structure-of-intellect model. Evaluation factors are concerned with abilities in which the intellectual process is reaching decisions or making choices about the goodness of information in terms of criteria of identity, consistency, and goal satisfaction. The evaluation abilities hold special importance to investigators of thinking in general because of the involvement of those abilities in almost all creative problem solving, particularly at the last stages of problem solving, where solutions must be accepted, rejected, or slated for revision. The symbolic factors are abilities in which the organism operates upon information in the form of signs or symbols such as letters and numbers.

The six factors selected for experimental investigation are: evaluation of symbolic units (ESU), evaluation of symbolic classes (ESC), evaluation of symbolic relations (ESR), evaluation of symbolic systems (ESS), evaluation of symbolic transformations (EST), and evaluation of symbolic implications (ESI). Although one of the above factors, ESR, was presumed to have been discovered (Guilford and Hoepfner, 1963), its existence as ESR was considered very tenuous and is therefore considered as an experimental factor in this study rather than a reference factor.

An objective of some importance in this study is that it represents another attempt to validate Guilford's model as a source of hypotheses for defining and isolating factors of human intelligence. If the model continues to generate concepts which can be found to represent unique abilities, then its contribution to psychological explanation and prediction will be further substantiated.

Not only is the existence of new factors deduced from the model, but the model further offers operational specifications for the measures needed for the factors. Such measures, which serve as the "empirical world" for verification of the factors and the model, are then available as instruments for study of the traits in new investigations involving those traits. Today's new factors and their experimental tests become tomorrow's reference concepts and marker tests for use in other kinds of investigations. They also become available for the applied psychological prediction and selection, which is the ultimate social-value testing of the model itself.

The potential value of the symbolic-evaluation tests as measures of constructs of intelligence is considerable over a variety of applied problems in psychology. Petersen, et al. (1963) reported that five of the newly designed symbolic - evaluation factor tests contributed to the prediction of success in high-school general mathematics and algebra. Three factors, ESR, ESS, and EST, each hypothesized to be represented by one test,

---

[1]This study is one of a series conducted by the Aptitudes Research Project at the University of Southern California, under Contract Nonr-228(20) with the Office of Naval Research, Personnel and Training Branch. The ideas expressed here are our own and in no way reflect the views of that agency. This material may be reproduced for any purposes of the United States Government. Among the authors, Guilford has served as Responsible Investigator and Director of the Project; Hoepfner as Assistant Director and Study Leader; Merrifield as Assistant Director during the early stages of this study.

contributed significantly to the predictions. The DAT subtest, Clerical, thought at that time to be a test of ESU (and the results of this study support that view), consistently contributed predictive variance, second only to the Numerical Ability subtest, when four DAT subtests were employed as predictors of success. There is considerable promise, therefore, that measures of symbolic-evaluation factors in general will aid in selecting or classifying students for mathematics programs.

Vinacke (1952) stresses the value of evaluative thinking, especially in the semantic area, as an important component in the final process of creative production, i.e., verification. There is no reason not to assume an equal value for symbolic evaluative thinking when the creative production is symbolic in nature. The creative mathematician and the mathematical-theoretical scientist must evaluate their numerical or symbolic formulations in terms of identity, consistency, and other criteria of goal satisfaction. In other words, each must employ his symbolic-evaluation abilities as he proceeds, and especially as one of the last steps in his creative production.

Guilford (1964) goes even further in describing his model of creative problem solving by stating:

> Another major feature (of the model) is the universal dependence of all operations upon evaluation. In general, behavior is self regulating and self correcting, through the principle of feedback information. An example is seen in the operation of evaluation, which helps to select information at the filtering stage near the point of input and to reject or accept information in the operations of cognition and production. In general, the operation of evaluation appears to serve a filtering function as well as other functions (p. 8).

Even in such current psychologically popular fields as computer simulation of thought processes, more exact knowledge of such abilities will be valuable. Attempts to program certain mental processes, particularly with regard to symbolic content, have repeatedly revealed the lack of information available on how human beings carry out seemingly simple thought processes. Such frustrated simulation research has led to the return to the study of human thought processes. The knowledge that is gained concerning the processes involved in the symbolic-evaluation abilities could very likely be useful in simulation research, which results, in turn, in greater knowledge of the thought processes involved.

REVIEW OF THE LITERATURE

Some General Sources

The philosophical literature is well-sprinkled with discourses on evaluation, choice, and decision making. In all cases, however, the major concern is with individual and cultural values and morality, and the experiential development of values. Empirical investigation of such concepts seems to be wholly lacking. These concepts would need a great deal of specific research before they could be employed as variables in a study like the present one.

Reports of educational research are primarily concerned with critical thinking, which appears to be limited to the evaluation of conceptual information in terms of logical criteria. The educator is interested in learning how to impart to his students the ability to think critically. Although such research is well worthwhile, it seems to be premature; the educator can hardly expect to teach critical-thinking skills when nothing but general and vague notions are held about their nature.

The psychological literature also abounds with reports on research bearing upon judgmental abilities. Depending upon the specific area of such research, different variables receive varying amounts of attention. In psychophysics, the major concern is with the relations between the stimuli and the subjects' judgments of them. The task in these psychophysical experiments is to find the lawful relationships existing between the continua of stimulus dimensions and the continua of judgment.

Attempting to develop a theory of judgment and choice, Restle (1961) states:

> A cognitive position on judgment argues that judgment depends both on external events (stimulus) and the observer, taken together as an interacting system (p. 189).

This position is clearly similar to that taken by the psychophysicists, but Restle represents part of a new development in the study of judgment—applications and explanations of judgment, choice, and decision through mathematical models. The probabilistic mathematical models have been applied to the investigation of decision processes under the subject areas of utility theory, information theory, decision theory, and game theory.

The present study is directly concerned with the very individual differences that the above-mentioned methods of investigation have ignored. The approach through individual differences is needed in order to derive the concepts descriptive of psychological events in decision making, making possible something more than a superficial approach to theory construction. Individual differences are generally investigated by means of mental tests and, currently, by factor analysis or some related multivariate procedure. For this reason, the review of the literature here will focus upon previous factor-analytic research.

Factor-Analytic Studies

Over the years, only by virtue of small, incidental steps was attention directed to the evaluation area by virtue of factor-analytic results. The first of these steps can be attributed to L. L. Thurstone (1938a) when he discovered the factor that became known as "perceptual speed," and that later was recognized as an evaluation factor—the evaluation of figural units (EFU). The factor has been most consistently defined by tests that require the rapid comparison of figural objects with judgments of identity versus non-identity (Guilford and Lacey, 1947). But there have been times when there was uncertainty as to whether it also applied to tests requiring the identification and matching of letters, numbers, and words (Thurstone, 1938b; Coombs, 1941).

Thus, the perceptual-speed factor is of double interest in the present investigation, since the structure-of-intellect (SI) model forecasts a distinct but parallel ability concerned with the identity versus non-identity of literal material, and tests of EFU could serve as a model for the hypothesized ESU factor, which heretofore had not been clearly demonstrated. One or two studies cited by French (1951) give some hope of such differentiation, e.g., Bechtoldt (1947), but there was nothing decisive.

Another step occurred in the Army Air Force Aviation Psychology Program, when a factor called "judgment" was found and verified (Guilford and Lacey, 1947). In earlier studies in the Aptitudes Research Project, however, such a factor did not at all consistently appear (Nihira, et al., 1964).

A factor that was found repeatedly in earlier Project analyses was first called "logical evaluation," most consistently brought out by tests of the syllogistic type. Since the tests were of multiple-choice form, alternative conclusions being given, not produced by the examinee, the interpretation was an evaluative ability rather than a deductive ability. There was a seemingly parallel factor, identified most consistently by a test called Symbol Manipulation, which was essentially syllogistic in form but with letter symbols rather than words. This is the nearest that the Project has come to demonstrating previously a symbolic-evaluation factor.

The Project's two previous investigations aimed specifically at evaluation abilities concentrated very much on the semantic factors. Hertzka, et al., (1954) confirmed the logical-evaluation factor, which was later interpreted as the EMR factor in the SI model. They also found a factor called "speed of evaluation," for the rather incidental (as it turned out) reason that it was a feature of highly speeded tests. A third evaluation factor was called "experiential evaluation," because it was distinguished by tests that called for judgments based on consistency with past experience. This factor was later placed in the cell for EMS in the SI model, since the consistency involved pertained to comparisons of parts of situations, which could well be regarded as systems.

In a recently completed Project investigation, the analysis was aimed at the six semantic-evaluation factors forecast by the SI model (Nihira, et al., 1964). The factor formerly called "speed of evaluation" was apparently verified, but it was defined by tests designed for factor EMU, hence was defined as the ability to evaluate semantic units. A new factor was found to represent the cell for EMC and one for the cell for EMR, and factor EMS was confirmed. The logical-evaluation factor, previously identified with the cell for EMR, was found to be better interpreted as factor EMI. The identification of five of the six semantic-evaluation factors lent much hope that a differentiation of six symbolic-evaluation abilities might be effected. There was also the possibility that to demonstrate a certain symbolic-evaluation factor, tests modeled after those for the parallel semantic factor might serve. Advantage was taken of such parallels in some instances, although the semantic-evaluation study was not completed before the symbolic-evaluation study was very well along, so that in most instances the model tests had not been established as good tests for their respective factors.

It is of interest to note that in a study designed to learn more about the intellectual abilities pertinent to success in mathematics, Canisia (1962) employed quite a number of symbolic tests, some of which appear to have some evaluation variance. Examples are: Number Series 2 (which would seem to be a measure of factor ESS); Conditions 2 (factor ESR); and Qualitative Relationships (factor ESI). Having only one representative test for each factor, however, one could not expect these three factors to emerge. But it would appear that Canisia was working in the right directions.

THE HYPOTHESIS

The formal hypothesis is concerned with the demonstration of six selected intellectual abilities predicted from the SI model, using the experimental tests designed to demonstrate them. Guilford (1960) has concisely stated the requirements of a good factor-analytic experiment whose objective is to demonstrate something of basic psychological significance. First, the investigator makes hypotheses, based upon reasonable inferences from previous research or upon implications from a model or theory, according to which individuals are expected to differ from one another. Such hypotheses are given attention in the first part of this section. Second, the investigator develops measures that will broadly sample behavior indicating the individual differences. The second part of this section is concerned with the measures developed to test the hypothesis.

Statement and Elaboration of the Hypothesis

The major hypothesis is that, in the population employed in this study, six separate factors of symbolic evaluation, predicted by the SI model, exist, distinct from one another and from factors represented in certain other areas of the model. Furthermore, these six factors are not restricted to any one kind of symbol. In other words, tests defining any one of the six factors could be composed of words, numbers, or letters. Moreover, each of the six factors is defined by tests involving both sensitivity and estimation, concepts to be explained later.

The formal hypothesis is somewhat complex, an attribute not conducive to unambiguous evaluation of results. It is necessary to look more closely at some of the specific aspects of the hypothesis. If one defines the object of this study as the evaluation of symbolic products, careful consideration must be given to each of the underscored words, in order to clarify all aspects of the formal hypothesis.

A Definition of Evaluation

In formulating a definition of "evaluation" for this study, several alternatives were considered. We shall mention two of them, since they have a direct bearing on the distinction between "sensitivity" tests and "estimation" tests, and then state a broader definition that embraces both conceptions.

In a definition that equates evaluation to "sensitivity to error," the term "error" is interpreted broadly to include any kinds of defects, deficiencies, departures, inconsistencies, incongruities, etc. This view implies absolute judgments; a thing is perfectly all right or is identical with another, or it is not. Some individuals can detect such "errors" where others cannot. Those who can are more able with respect to evaluative abilities than those who cannot. This is not to say that there is a dichotomy of individuals; they can still vary by small degrees along a continuum of greater or less sensitivity.

In the definition that emphasizes "estimation," it is implied that individuals also make relative judgments. When items of information fall short or deviate from a standard, one may deviate farther than another. It may be obvious that all the items of information depart from the standard, but which one deviates least? Where sensitivity tests typically call for absolute judgments of a yes-no, disjunctive type,

estimation tests typically offer alternative items of information and ask which one deviates least, or sometimes (but rarely) which one deviates most. A ranking of alternatives is implied. This view of estimation is concordant with that offered by D. M. Johnson (1955) when he says,

> When judgments are expressed on a continuous scale, . . . the term estimating may be used, as a special case of judgment. Because the scales used in estimation are continuous, there are not boundaries between the categories . . . (p. 284).

Actually, the two views can be brought logically under the same definition of evaluation. In both cases, a standard of some kind is implied. In both cases, criteria for judgment are implied. A definition embracing both views would read: "Evaluation is a matter of decision concerning criterion satisfaction." This is the definition adopted as the basis for planning this study. One of the objectives of this study was to determine whether tests embodying the sensitivity principle and those embodying the estimation principle would both indicate the same kinds of ability, thus justifying subsuming them both under a single definition of evaluation.

Criteria for Evaluation. When it is said that evaluation is concerned with criterion satisfaction, it is then necessary to give attention to what kinds of criteria are suitable for use in tests of evaluation abilities. Some of the traditional criteria have been: identity vs. deviation, completeness vs. incompleteness, compatibility vs. incompatibility, congruity vs. incongruity, effectiveness vs. ineffectiveness, and suitability vs. unsuitability. For testing purposes, a criterion must be of a type that can be communicated to the examinee. As will be seen in discussions of the results, some additional kinds of criteria were included in this study, such as popularity vs. unpopularity (frequency of usage) and highly probable vs. improbable. With symbolic material to be evaluated, there was no question about using either esthetic or moral criteria. Questions regarding those two kinds of criteria will eventually arise in analyses pertaining to figural information on the one hand and behavioral information on the other.

Definition of Symbolic Information

The whole logic of symbolic communication as compared with conceptual or semantic communication is that more precision can be had due to the denotative inflexibility of symbols. One might then ask the question, "What is there to evaluate in connection with symbolic information?" We might expect some different aspects to evaluation of symbolic information than those that apply to semantic information, which is relatively rich with connotative meaning.

Several different varieties of information conform to the definition of symbols stated by Guilford and Hoepfner (1963): "Information in the form of signs, having no significance in and of themselves." (p. 2). The clearest example of a symbol is a number. Numbers have no significance in and of themselves, yet can be evaluated for numerical identity, order, or consistency, with respect to other numbers. Letters also conform to the definition when they are processed in terms of their literal properties rather than their figural properties. Syllables can be symbolic units, as well as words, when their semantic meanings are not relevant to the task, as in breaking words into syllables or in word compounding. All these types of symbols were used in various experimental tests in this study.

Definition of Products

The six kinds of products of the SI model have been defined in a number of places, but will be very succinctly defined here. Units are segregated items of information having thing character. Classes are groups of items of information having common properties. Relations are meaningful connections between units. Systems are organized or structured complexes of units and relations. Transformations are changes or redefinitions of known items of information. Implications are in the form of expectancies, predictions, or consequences of information.

This aspect of the hypothesis, the existence of six distinct product factors of symbolic evaluation, provides the major problem of this study. The demonstration of these six factors, or the failure to do so, would confirm or fail to confirm the model from which the hypothesis was deduced. The tests designed as measures for each of the products will be described in detail later.

Varieties of Experimental Tests

With three kinds of symbols available and with the distinction between sensitivity and estimation tests, for a completely systematic experimental design it would have been desirable to have six experimental tests for every product factor. No effort was made to achieve fully this kind of coverage with experimental tests, and it was difficult to achieve all six kinds of tests with each product. Table 1 shows how far the test development went in the direction of complete, systematic coverage.

The two major columns represent sensitivity vs. estimation types of tests; those involving absolute judgments vs. those requiring relative judgments. Comparing the two columns it will be seen that the balance is fair. For particular factors the balance is good in the case of three, but not balanced in the other three. Overall factors there is a fair balance between the three kinds of symbols, with a few tests being composed of combinations of two or three kinds. There proved to be enough dispersion of the conditions to make possible answers as to whether both sensitivity tests and estimation tests serve to measure these evaluation factors and whether the kind of symbol makes a difference in the success of tests.

Descriptions of Factor Measures

Reference Factors

An important goal for the well-designed factor-analytic test of factor hypotheses is not only to determine what the experimental constructs are but also what they are not. The formal hypothesis states that the six factors of symbolic evaluation are not only distinct from one another, but are also distinct from other factors deduced from the model. For this reason, a number of marker tests, known from previous experience to measure reference factors, were included in the analysis to demonstrate the distinctness of the new experimental factors from factors already known.

Table 1

Semi-Balanced Factorial-Like Design of 25 New Symbolic-Evaluation Tests

Product	Format	Sensitivity	Estimation
Units	Numbers		
	Letters	Letter "U"	Familiar Letter Combinations
	Words	Correct Spelling Derivations	
	Combinations	Symbol Identities	
Classes	Numbers	Best Number Pairs	Best Number Class
	Letters		
	Words	Sound Grouping	Word Choice
	Combinations		
Relations	Numbers	Sign Changes II	
	Letters	Symbol Manipulation	
	Words	Similar Pairs	Related Words
	Combinations		
Systems	Numbers	Correct Number Series	Series Relations Way-Out Numbers
	Letters	Correct Letter Orders	Best Letter Set
	Words		
	Combinations		
Transformations	Numbers		
	Letters		Typing Errors
	Words	Jumbled Words	Decoding
	Combinations		
Implications	Numbers		
	Letters		Letter Problems
	Words		Abbreviations
	Combinations	S Test	Symbol Reasoning

The usual test of the distinctness of experimental factors is made by selecting for simultaneous analysis reference factors that might possibly be identical with the experimental factors. Of all the non-evaluation factors, those of cognition were suspected of being most likely to be confused with the experimental evaluation factors. One reason is that it takes care to construct an evaluation test so that it does not offer necessary cognition problems of sufficient difficulty to introduce some cognition variance into the total scores, or, indeed, that it does not become a cognition test instead of an evaluation test. For this reason, tests of five parallel symbolic-cognition factors were selected for inclusion in the battery.

A measure of numerical facility, recognized as a memory factor in the SI model, was included to define that factor. The reason was the large proportion of number tests in the battery, some of which might have some numerical-facility variance that should be segregated.

Because scores on one of the classes-evaluation tests (Best Number Class), designed for the factor ESC, were expected to depend upon ability to produce rapidly responses to be evaluated, two measures of the parallel divergent-production factor, DSC, were employed as marker tests. In addition, this analysis afforded an opportunity to learn more about the nature of the DSC factor, which had not previously been very clearly demonstrated.

It might be expected that most convergent-production tests would involve some evaluative activity, since competing answers to items must be rejected in favor of the one right answer, hence such tests might have some evaluation variances. Three of the well-known symbolic convergent-production factors were represented by two marker tests each.

To determine clearly the separateness of the perceptual-speed factor EFU from its parallel symbolic counterpart (ESU), two marker tests for EFU were included. This was expected to clear up the ambiguity that had been left with regard to these two factors in previous investigations, as discussed earlier.

Although not considered as a reference factor, an anticipated motivational variance seemed to need accounting for. It was expected that symbolic tests might have systematically different degrees of appeal for individuals, which might contribute to the appearance of a general evaluation factor or a substantial appearance of obliqueness among the evaluation factors. In order to segregate such sources of variance, four experimental measures were included in an attempt to isolate a possible motivation factor.

The Marker Tests

In the following paragraphs, each of the 13 reference factors and their marker tests is mentioned in turn. Descriptions of the tests, with sample items, will be found in the Appendix of this Report. Most of the symbolic marker tests had been used in a previous analysis (Guilford, et al., 1961). Test data from the testing in that analysis were used in item analyses that were aimed toward shortening the tests by eliminating items that (1) did not correlate well with total scores of their own factors and (2) correlated with total scores of other factors. In addition to saving testing time, this procedure was expected to make the total scores more nearly univocal for their own respective factors.

CSU Tests

Disemvowelled Words—E recognizes words with vowels omitted.

Word Combinations—E sees what word can be made out of the end of one word and the beginning of the next.

CSC Tests

Number Classification—E selects a number that goes in a class with three other numbers.

Number Group Naming—E states how three numbers are alike.

CSR Tests

Seeing Trends II—E identifies and names the trend in a series of words, based upon alphabetical principles.

Word Relations—An analogies test in which spelling relations between words must be seen.

CSS Tests

Circle Reasoning—E sees a principle of order in series of circles and dashes.

Letter Triangle—E sees the alphabetical principle by which letters are arranged in a triangle.

CSI Tests

Symbol Grouping—E rearranges a line of scrambled X's, O's, and -'s in systematic order, making as few moves as possible, involving foresight.

Word Patterns—E arranges words in cross-word-puzzle fashion, as efficiently as possible.

CMU Tests

General Vocabulary—Test 8 of the Iowa Tests of Educational Development (ITED) (Lindquist, 1959).

Verbal score of the Preliminary Scholastic Aptitude Test (PSAT) (ETS, 1959).

Verbal score of the Cooperative School and College Ability Test (SCAT) (ETS, 1958).

MSI Tests

Numerical Operations—Part III of the Guilford-Zimmerman Aptitude Survey (1956).

DSC Tests

Number Grouping—Given a set of several numbers, E arranges and rearranges them in different classes.

Varied Symbols—E classifies letter sets in different ways.

NSS Tests

Operations Sequence—E states the correct order in which a series of numerical operations must be performed.

Word Changes—E gives the order in which short words must be changed, one letter at a time, to go from starting word to goal word.

NST Tests

Camouflaged Words—E breaks up consecutive words to find names of sports or games.

Word Transformations—E regroups a series of words in a phrase to obtain a new phrase.

NSI Tests

Form Reasoning—E solves equations composed of figure-symbol combinations.

Sign Changes—E solves simple numerical equations, following rules about interchanging signs.

EFU Tests

Identical Forms—E selects figures identical with a standard.

Perceptual Speed—Part IV of the Guilford-Zimmerman Aptitude Survey (1956).

Experimental Factors

The tests for the hypothesized factors were developed using either of two approaches or a combination of the two.[2] In the first approach, specific examples of tasks are deduced from the operation-content-product combination being investigated. For example, the ability in the cell EST, evaluation of symbolic transformations, involves evaluation of changes in symbolic materials. A code can be an example of a symbolic change, and so the test Decoding was developed.

The second approach emphasizes tasks similar to those for established factors having one or two attributes in common with the new factors. For example, ESU, evaluation of symbolic units, and EFU, evaluation of figural units, differ only with respect to the content category; test formats might be very similar. Four of the 25 symbolic-evaluation tests were revised from older tests. They are Letter "U", Sign Changes II, Symbol Manipulation, and Sound Grouping. The remaining 21 tests, developed by the first and second approaches, were written by the first author.

During the period of September, 1961, through March, 1963, nearly 30 different pretest booklets were administered to classes in psychology at the University of Southern California, Los Angeles City College, and Pasadena City College.[3] These pretestings were designed to obtain technical information such as the appropriate level of item difficulties, comprehension level of the test instructions, test reliabilities, and optimal time requirements for newly developed tests. Extensive item analyses were conducted whenever pretesting information revealed low reliability estimates.

From the reliability and intercorrelation data obtained from pretesting, 25 tests were selected from a pool of over 40 tests especially designed or adapted to measure the six experimental factors. The selected tests had pretest reliabilities in the 70's and .80's. Within-factor intercorrelations were generally considerably higher than between-factor intercorrelations, further ensuring the demonstration of the distinctness of the expected factors. The criteria of high reliability and desirable correlational pattern, in conjunction with the considerations listed in Table 1, determined which tests were finally selected to represent the experimental factors in the final analysis. In the following paragraphs, the six experimental factors and the tests selected to measure them are discussed in detail.

ESU Tests

The five experimental tests developed to measure evaluation of the product of units had in common symbolic stimuli that are processed as wholes, rather than separated, analyzed, or classed. Although similar symbolic stimuli are employed in tests of the other intellectual products, the mental process performed upon units must maintain the thing quality of the stimuli. Previous tests that have aided in defining factors of symbolic units involved the following kinds of stimuli: anagrams, words, scrambled words, and incomplete words (CSU); digits and letters (MSU); specified classes of words in terms of given initial letters, suffixes, and prefixes (DSU). Guilford and Hoepfner (1963) suggest tests employing letters and digits as stimuli for measures of ESU based on tenuous prior studies (French, 1951). Construction of the experimental ESU tests was based upon the history of parallel tests of symbolic units and the try-out of new kinds of stimuli.

The test Correct Spelling employed complete, common English words as symbolic stimuli. The words function as symbols because E is to direct evaluation toward spelling rather than meaning. E is tested on his sensitivity to the correctness or incorrectness of the spelled symbolic unit. In this case, sensitivity to spelling is based largely upon the long-term retention of the correct symbolic elements of standard English words. The words employed as items were selected from lists of commonly misspelled words published in English handbooks and secretarial manuals. The test was designed as a power test, there being enough time to judge all spellings. Tests very similar to Correct Spelling were found by Thurstone to be loaded on the word-fluency factor (1938a) and on a verbal factor (1940).

Derived from the format of the test Anagrams, used by Thurstone (1938a), Derivations also employs complete English words as test stimuli. Whereas Thurstone's test has Es make as many short words as they can in a limited time from the letters in a large given word, Derivations supplies not only the given word, but also 50 short words derived from it. E's judgments are based upon sensitivity to the errors in some words that could not be derived from the long given words. The task of this test, detecting changes in anagrammatically-derived words, naturally suggests that the product to be evaluated might be transformations. Although Derivations was originally written as an EST test and there may be some EST variance involved in its scores, pretesting led the authors to hypothesize that the scores would be more heavily determined by ESU variance. The reason probably is that it is the end product of a change that is to be evaluated, not the change, as such. A similar problem was encountered in constructing tests for EMT (Nihira, et al., 1964).

The given stimulus words were selected on the basis of the number of different letters they contain: PARENTHESIS, REPUTATION, and COMBINATION. Errors of some derived words involve the use of letters not in the given word or the use of a letter more

---

[2] Much credit is due Miss Carolyn Frame, who assisted in the construction and preliminary analysis of the experimental tests.

[3] For much of the pretestings we are indebted to Dr. Constance D. Lovell of the Department of Psychology, U. S. C., Mr. R. Jones, U.S.C. Testing Bureau, and Mr. H. Petersen, Pasadena City College.

often than it appeared in the given word. Derivations was designed as a speed test. The instructions urge the Es to make their judgments rapidly as there would be insufficient time for careful consideration of every derived word.

Familiar Letter Combinations is an experimental test that has a completely new type of symbolic stimuli: three-letter syllables. E is to estimate which of two given syllables is more common as a part of real English words. Familiarity is the criterion for decision. Neither the syllables nor the criteria of real words are to be considered semantically; only the relative frequencies of occurrence are relevant. The key for this test was determined from the empirical frequency counts reported by Underwood and Schulz (1960). The nonsense syllables are paired so that the keyed syllable is far more commonly used than its alternative. Because of its exceedingly experimental and evaluative nature, Familiar Letter Combinations is timed as a power test.

Although the highest reliability estimate determined from pretestings was low, and correlations with other tests of ESU were small and even negative, Familiar Letter Combinations was used as a test of ESU in the final testing because of its novel content and its apparent fit to the ESU hypothesis.

Letter "U" is a test of E's sensitivity to the presence of a specified letter in words under speeded conditions. It is based upon Thurstone's test Letter "A" (1938b), which split its variance among factors that Thurstone called perceptual, number, and word factor. Bechtoldt (1947) found Letter "A" to be loaded highly on a factor with a test of crossing-out specific letters on a page of regularly spaced letters. He named the factor: "speed of recognition of predetermined symbols in context of discrete distractors." Although Cattell names the factor on which this test is loaded "speed of symbol discrimination" (Cattell, 1953), and Guilford and Hoepfner (1963) suggest the factor is ESU, French, et al. (1963) conclude that our knowledge concerning this factor is not at all clear since several "sub-factors" tend to pull together in different ways, depending upon the tests included in the factor-analytic battery. In general, a test like Letter "U" is often found in strong relation to perceptual-speed tests. To clarify this ambiguity, not only were four tests designed to measure ESU included in the battery along with Letter "U", but also two strong perceptual-speed (EFU) tests.

The test Symbol Identities was designed as a measure of E's sensitivity to the identity or non-identity of paired sets of numbers, letters, and words, under speeded conditions. It is essentially parallel to tests of EFU, in which identity of pairs of figures is in question. Symbol Identities is the only ESU test employed that cuts across all the possible stimuli considered appropriate in the symbolic domain.

Symbol Identities is similar to many of the tests designed to measure clerical speed and accuracy; E decides whether or not the two members of pairs of symbol sets are the same or different. This test, like Letter "U", could conceivably share much figural variance, as Es could compare each symbol stimulus, figure by figure, and arrive at an accurate judgment. Such activity is very inefficient, however; a figural attack upon Symbol Identities should result in poor performance, unless it is used only when a quick symbolic attack does not yield a decisive choice.

ESC Tests

Four experimental tests were developed to measure the factor of evaluation of symbolic classes. A symbolic class was defined for this investigation as a group of symbols with some common property. Such a group of symbols would be composed of at least two members whose common property must be symbolic, not figural or easily semanticized. Two of the four tests designed to measure ESC were symbolic adaptations of tests of semantic evaluation of classes; another test was rewritten from a test that had no firm construct "home"; and the remaining test was constructed originally for this investigation. Tests of symbolic classes previously employed such stimuli as number groupings (CSC and DSC) and number pairs (CSC) (Guilford and Hoepfner, 1963).

Best Number Class, the new test developed for this study, has no history in previous studies of intelligence. E's task is to assign given numbers into one of four classes in such a way as to maximize each number's value by assigning it to the most exclusive class it fits. The four classes into which the stimuli were to be assigned were, in order of exclusiveness: EVEN MULTIPLES, ODD MULTIPLES, SQUARES, and PRIMES. Es were warned that the numbers could possibly be assigned to several classes and that credit could only be earned by assigning each number to its most exclusive class.

The test Best Number Pairs is the other hypothesized ESC measure employing numbers as stimuli. The item format is adapted from the test Best Word Pairs, found to be loaded, in part, on the factor EMC, evaluation of semantic classes (Nihira, et al., 1964). In the symbolic form, E's task is to choose one of three pairs of numbers that makes the best class. In order from best to poorest, the classes are: pairs of perfect squares, pairs of multiples of the same number, pairs of odd or even numbers, and pairs with no class property. Although the number classes in this test are similar to the classes in the test Best Number Class, the task in this test is close to a sensitivity-type of test, whereas the task called for in Best Number Class is close to the estimation type, based upon an understanding and memory of the number classes. Like its semantic counterpart, Best Number Pairs was developed as a power test.

Sound Grouping is a test with a long history. In each item, four words are given, three of which are fairly good rhymes and the fourth is not. The latter is to be noted and selected, for the right answer. The test's factorial composition has been open to considerable question because of its tendency to go with different factors, depending upon the battery in which it has been analyzed. Thurstone found that its variance split over his factors of word fluency, spatial, and verbal relations (1938a). Later, the test was employed in research at the University of Southern California, and was variously found to split its variance over the factors of verbal comprehension (CMU) and general reasoning (CMS) (Green, et al., 1953); verbal comprehension and eduction of correlates (NSR) (Guilford, et al., 1954); and eduction of correlates, facility with verbal relations, and perceptual speed (EFU) (Hertzka, et al., 1954).

Because of this history of factor instability and the fact that the above-mentioned studies did not include tests of what would now be called symbolic classes, Sound Grouping was hypothesized to measure ESC, a seemingly logical place for it. It was not expected, however, that Sound Grouping would suddenly become a uni-factor test when placed in a battery with several ESC tests.

There is some evidence that Sound Grouping may measure some figural ability, as Hertzka, et al. (1954) suggest in their explanation of its loading on the perceptual-speed factor. It would be in an auditory-units category, however, rather than visual. The four word stimuli for each item must be processed as auditory figural units before the word with the different sound can be determined. Once the figural stimuli are before E, the task is simply one of cognition of the stimulus that does not belong. The symbolic-evaluation process probably enters at the stage of translating the visual symbolic and semantic information into auditory figural information, based upon the varying sounds of letters in combination. It was hypothesized that the ability to make the translations rapidly is facilitated by symbolic sensitivity to auditory figural correlates of letter combinations. This detailed logical analysis resulted in the hypothesis that Sound Grouping might share considerable ESC variance. Several pretestings confirmed the idea in that Sound Grouping did correlate slightly with other ESC tests.

The fourth test designed to measure ESC is Word Choice. The item format was adapted from the semantic classes test, Class Name Selection (EMC), which was developed by the first author and which aided in the definition of the EMC factor in the study by Nihira, et al. (1964). Whereas in the semantic test E is to choose the best class name for a group of four words, in the symbolic test E is to choose the best of three possible additions to a class of three words. The class properties used in Word Choice are symbolic, e.g., order or nearness of certain letters or types of letters in the words. This test also differed from its semantic counterpart in that none of the alternative words for any class completely possessed all the class properties; a best word had to be chosen, even though it was slightly wrong. It is thus an estimation test.

ESR Tests

The four tests developed to measure the factor ESR, evaluation of symbolic relations, employed recognized connections, based upon symbolic variables, between symbolic units. Examples of connections based upon symbolic variables are "bigger than," "equal number of consonants," and "similar ratios." Although ordering tasks in the symbolic area appear often to be more relational than systemic, considerable difficulty in developing evaluative forms of the task precluded development of such tests to parallel more closely the tests of other symbolic-relations factors. It must be remembered, that as with the product of classes, the perceptual stimuli themselves need not be symbolic, but the connections between the stimuli must be based upon some symbolic aspect of the stimuli.

In the divergent-production domain of thinking processes, Gershon, et al. (1963) states that it is easier to invent relational tasks that require the introduction of alternative relationships among symbols than it is to invent completion-production items as are so frequently seen in the semantic tests of relations. Although completion-production formats are not typically relevant to tests of evaluation abilities, the task restrictions in divergent-production processes are different from those for evaluative processes. Two of the experimental ESR tests introduce alternative relationships, while the other two involve evaluation of a relation in terms of provided rules.

Stimuli employed in previous symbolic-relations tests have included: word trends and alphabetical relationships (CSR); number pairs and word pairs (MSR); letter groups and numerical equations (DSR); word pairs and letter series (NSR); and equations (ESR) (Guilford and Hoepfner, 1963). In addition to employing numerical equations as test stimuli for ESR tests, words were employed in two tests, so that the four tests were equally divided between number-letter and word stimuli.

The first experimental ESR test, Related Words I, was adapted by analogy to Matched Verbal Relations, designed for factor EMR, evaluation of semantic relations, on which it was loaded in two studies (Nihira, et al., 1964; Petersen, et al., 1963). In Related Words I, E must estimate which of three alternative word pairs is most similar to a given related pair. The relation between members of any pair is based upon the order and position of letters and the vowels and consonants that are changed or moved. This is the only ESR test in which no alternative answer is completely correct; only a best alternative is to be selected. Like the other ESR tests, Related Words I was designed to be a power test.

Sign Changes II had been developed as an ESR test to be used in a predictive battery for success in ninth-grade mathematics courses (Petersen, et al., 1963). It was not factor analyzed, but multiple-prediction studies did indicate that the test contributed unique variance to prediction of success to a significant degree. The task in this test is to determine what sign changes, if any, must be made to change a numerical expression into an equation. An elementary understanding of arithmetical operations and the relationships of equality and inequality of expressions is all E needs in order to understand clearly the test items and the task. Because, for each item, there is a unique, correct answer, and because the test is not speeded, for Es well-versed in such algebraic-equation manipulating the evaluation variance might be minimized in favor of variance determined by individual differences in a convergent-production activity. Results of the Petersen, et al., (1963) study, however, showed that the correlations between Sign Changes II and convergent-production tests were quite low. It was hypothesized that even at a somewhat advanced stage of algebraic competence (in an accelerated-algbera group), a sensitivity to the correct relationships is at the basis of the process necessary to perform on this test.

Similar Pairs is a new test, in both idea and items. The stimuli are word pairs, the members of which are related by letter locations and letter changes. E's task is to judge whether the members in two such pairs are or are not similarly related. The process involved in responding to this test is sensitivity to sameness or differentness of the relations within the word pairs. In all the items, the relations within the pairs were kept extremely simple, so that there would be little or no difficulty in cognizing the relationships, so that cognition variance would be minimized in the test scores and thereby maximize the relative importance of the evaluation variance.

Symbol Manipulation is a test of the ability to decide whether a given relationship between two letters follows logically from other statements of relationship involving the same letters, where the relationships are "greater than," "equal to," and "less than," and their negations, all statements in symbolic form. This test has appeared in five previous analyses done by the Aptitudes Project. In two analyses it had no significant loadings on any factors, probably because there was a paucity of symbolic tests and, more particularly, because the right symbolic factor was not determined (Berger, et al., 1957; Wilson, et al., 1954). In the first analysis in which the test had significant loadings (Green, et al., 1953), they were on a factor called "logical reasoning" (later, logical evaluation, and eventually identified as EMR) and on

a factor called "symbol substitution," which might have been the same as one later called "symbolic manipulation" for the test of the same name. In the second analysis (Guilford, et al., 1954), Symbol Substitution again shared its variance with the logical-evaluation and symbol-manipulation factors. In a third analysis (Kettner, et al., 1959), Symbol Manipulation again led uniquely on a factor given the same name. In the same analysis a symbol-substitution factor was found separated from symbol manipulation, which lends some support to a conclusion stated above about an earlier symbol-substitution factor. It was the parallel nature of Symbol Manipulation and verbal-syllogistic tests of logical evaluation, recognized as EMR, and the fact that this test also showed some variance in EMR (probably due to E's verbalizing the statements) that led Guilford and Hoepfner (1963) to recommend the test for ESR.

ESS Tests

Like the tests hypothesized to measure ESU, tests for ESS, evaluation of symbolic systems, seemed to be easy to construct. Almost one dozen tests were developed to measure ESS and were pretested. Most of the tests at this experimental stage proved to have reasonably good reliability and reasonable intra-factor correlations. The five tests chosen to define the systems factor in the final analysis broadly cover the various types of symbolic content and sensitivity vs. estimation.

All the stimuli for tests of symbolic systems are organized aggregates of units or relations wherein the interrelated or interacting parts are symbolically defined within the aggregate. The system, then, is the organization or pattern of parts, which may be compared with another system as to identity or similarity or evaluated for internal consistency. Other tests known to measure factors of symbolic systems have employed: sequences and alphabetical patterns (CSS); number-letter codes and equations (DSS); and sequences of numerical operations and of letter operations (NSS) (Guilford and Hoepfner, 1963).

Best Letter Set was designed as a measure of E's ability to estimate which of three sets of three or four letters each is most like a given set. The criterion of similarity is based upon the order and kinds of letters within the set. Although such small sets of letters might appear to function as units, the systems qualities of the alternative sets were sufficiently similar to force E to focus on them. It seemed highly unlikely that even the most sophisticated E could treat the stimuli as units and obtain a high score on this test.

The next two tests are somewhat alike and will be discussed together. Both Correct Letter Orders and Correct Number Series are tests of E's sensitivity to internal inconsistencies in symbolic systems. The stimuli in both tests are sequences of symbols organized according to some simple principle, similar to items in familiar number-series tests. The systematic order is stated verbally and E is to judge whether or not the sequence follows that principle. This test uses as stimuli, sequences like those used by Thurstone (1938a, 1940). Thurstone's tests were of the completion variety that emphasize the cognition aspect of the task; E cognizing the systematic sequence and showing that it is cognized by continuing the sequence in a consistent manner. Thurstone's tests, in both letter and number forms, were loaded on factors called deductive and inductive. More recently, series-type tests have been found loaded on the factors of general reasoning (CMS), eduction of patterns (CSS), and symbol manipulation (Kettner, et al., 1959); and on a factor called NSR (Petersen, et al., 1963), which was defined in part by tests of CSS and NSS.

The test Series Relations might also be considered an evaluative form of a number-series test, even though the task appears to be quite unlike that for Correct Number Series. In Series Relations, E is given a series of three numbers and is told that each element of the series except the first one is determined from the previous element (one to the left) according to some unknown rule. E is then to estimate which of three alternative rules or operations would best relate each series element to the previous one. Although E might simply try each rule upon the first and second series elements, obtain a three-number series, and compare it to the given one, selecting the correct rule, he is forced into making a choice or judgment because none of the three alternative rules is fully correct. That is, no one rule will correctly reconstruct the series from the first element; but one will do the job best.

In the test Way-Out Numbers, E is presented with a list of four ordered numbers and is instructed to choose either the first or last one on the basis of its being farther away from the remaining three numbers. In other words, E is to arrange the numbers on the dimension of numerical value and is to choose that extreme number whose value is farther from the other numbers' values.

E is provided with an example of a graphic solution in the instructions and is also told that there are mathematical methods of solution (compute the centroid of the first three numbers and determine its distance from the last, then do the same thing for the last three numbers, comparing the two distances). To force E into making rapid estimations, he is told that either procedure, while yielding correct answers, would be inefficient in terms of the time spent on each item. There was no evidence from the pretesting results that Es ignored this advice.

All the stimuli used in this test formed highly unevenly distributed series, which should make a number-estimation task possible. After the test had been printed, however, it was discovered that there was an easier method than roughly estimating in order to make correct responses. Owing to lack of foresight and consequent item revision, it was discovered that the correct responses could be chosen simply by subtracting the extreme numbers from their adjacent numbers. The larger difference, then, indicated the number farther away. If many Es discovered this principle and employed it, in spite of instructions, it would be expected that Way-Out Numbers would share its variance with the numerical-facility factor (MSI), due to the contribution of correct subtraction.

EST Tests

Three experimental tests were selected to measure the factor EST, evaluation of symbolic transformations. The extreme difficulty of constructing reliable evaluative tests of symbolic transformations limited the choice of tests. It was expected, however, that the test Derivations, hypothesized as a measure of ESU, might aid in defining the transformations factor. The content of the three tests was concerned with changes from one form of symbol to another equivalent form, or changes in symbolic units to meet certain requirements. The transformations tests developed for this study used letters and words as stimuli. It appeared, during test construction, that numerical stimuli were not readily susceptible to transformations without the involvement of other products, such as relations or systems. The denotative inflexibility of numbers did not allow for equivalent forms of the same numerical value using two different symbols. This limitation also obtained in the study isolating the only other known symbolic-transformation

factor, NST (Guilford, et al., 1961). The tests loading on the NST factor all involved words as the stimuli.

A rather common transformation of words and letters is any code that allows their encoding. In almost all cases of symbol coding, the transformation of one set of symbols to its encoded set of symbols is a one-to-one mapping of the symbol set onto the code set. Such a one-to-one mapping is suitable for a test of sensitivity to errors in coding only when the test is speeded and the coding system is well known by the Es. This implies that the sensitivity to slight, but possibly important miscodings is an evaluative process for the individual who functions well (is experienced) in the coding process.

Because no coding system is known with great generality within the population, and because it is inefficient and self-defeating to teach Es a complete coding system (memory factors might predominate), the EST test, Decoding, employs a simple and ambiguous code. Simplicity and ambiguity were introduced into the coding system for Decoding by employing a code for letters which does not map one-to-one onto the alphabet. The ambiguity of the code allows for words to be judged according to their ease of encoding or decoding. The change from an unambiguous code to an ambiguous one also changes the type of evaluation test involved. Whereas an unambiguous code and experience call for sensitivity, an ambiguous code calls for estimation; the code provides incomplete information, and E must estimate the complete information.

In the test Decoding, E is presented with two words and is asked to choose which one, if coded, would be easier to decode unambiguously. E is also given the opportunity to judge both words as equal in difficulty of decoding.

Jumbled Words is the only test designed for EST that is in the sensitivity category. E is given a stimulus word containing between five and seven letters and is to judge whether or not each of five alternative words is an accurate anagrammatic derivation from the given word. Jumbled Words is, therefore, similar in stimulus material to the test Derivations, which also uses anagram-type stimuli. Pretesting indicated that Jumbled Words correlated higher with EST tests than with Derivations. It was thought that this correlational pattern was due to the fact that Jumbled Words employed all the letters of each word, so that the whole word had to be evaluated for errors in transformation. This criterion is not present in Derivations, wherein only the same letters as those in the given word were the criterion, and fewer could be used.

The third test designed for EST, Typing Errors, is similar to Decoding in the task involved and the stimuli used. E is given an incorrectly typed word and is to choose from among alternatives the word that the incorrectly typed word would most likely be. The judgments are made on the basis of common typing errors due to the arrangement of the typewriter keyboard. A keyboard diagram is printed on each test page for E's reference.

The estimation process involved in responding to Typing Errors is probably not dependent upon EST ability alone, however. It would seem that some figural ability would be involved in this test due to the spatial nature of the keyboard arrangement. Further, it might be expected that typing experience might enter into proficiency at the required task. To determine to some degree the relative contributions of these two extraneous influences, it was decided to include at least one figural factor in the factor analysis and also to correlate test scores with a measure of typing experience.

ESI Tests

Tests designed to measure the factor ESI, evaluation of symbolic implications, employed all types of symbolic stimuli in the list. For the evaluation process, implications are defined as the expectancies or probable relative values of the presented symbols (estimation), or possible symbolic interpretations of a unit or system (sensitivity to symbolic problems). Previous investigations of symbolic-implications factors have employed the following stimuli in the tests: crossword puzzles and disordered symbols (CSI); simple numerical problems (MSI); words and equations (DSI); and simple equations (NSI) (Guilford and Hoepfner, 1963). The ESI tests developed for this study were based upon the stimuli used successfully previously and upon novel stimuli that appeared appropriate for evaluative processing.

The test Abbreviations presented E with a shortened spelling of a common word, E to choose one of the three alternative words that the abbreviated word most likely implies. The meanings of the words are irrelevant to choosing an alternative, and the spelling of the alternatives is correct. The only task for E is to choose the most expected value for the abbreviation, a task of estimating. No observance of shorthand principles was exercised in test construction; the abbreviations were short and relatively unambiguous. Usually, but not always, this implied dropping vowels and unsounded consonants from the keyed alternative. E was warned, however, that sounding-out the abbreviation would not necessarily aid him in his choice.

Abbreviations had been employed in a previous study at the University of Southern California (Petersen, et al., 1963), but was not factor analyzed. Its highest correlations in the two analyses reported were with the test Correlate Completion II, a test of factor NSR. Its reliability was reported as .47. Abbreviations was item-analyzed and lengthened for this investigation in order to improve its reliability.

Letter Problems is similar in format to Form Reasoning, a test of NSI. In Form Reasoning, E makes appropriate substitutions of symbols and solves a simple equation. The evaluation form uses letters as the stimuli and asks E not to solve the equation, but to judge the difficulty or possibility of solving it, on the basis of provided rules. It was hypothesized that E would have to make his judgments based on foresight. E's judgments were of the three-category type; problems were easy to solve (straightforward), difficult to solve (involving manipulations), or impossible to solve due to inadequacies of the table of substitutions.

The third ESI test is named S Test. The "S" in the name means that the test is symbolic. In this test E is given a stimulus about which he is to find a problem to solve. The solution indicates the nature of the problem to which E was sensitive.

The S Test is an adaptation of the F Test (Wilson, et al., 1954), designed as a test of sensitivity to problems, which had its highest (but insignificant) loading on the originality (DMT) factor instead. Although its behavior was not according to hypothesis, it was decided to revise the test to lead E into symbolic responses in all items; some were previously verbal. Then the test might measure E's sensitivity to symbolic implications of unstructured problems. It should be noted, however, that this test is not congruent with the conception that evaluation is "sensitivity to error," for no error is judged. It is a test of E's sensitivity to implications (as, indeed, it turned out) rather than a sensitivity to errors in implications.

factor, NST (Guilford, et al., 1961). The tests loading on the NST factor all involved words as the stimuli.

A rather common transformation of words and letters is any code that allows their encoding. In almost all cases of symbol coding, the transformation of one set of symbols to its encoded set of symbols is a one-to-one mapping of the symbol set onto the code set. Such a one-to-one mapping is suitable for a test of sensitivity to errors in coding only when the test is speeded and the coding system is well known by the Es. This implies that the sensitivity to slight, but possibly important miscodings is an evaluative process for the individual who functions well (is experienced) in the coding process.

Because no coding system is known with great generality within the population, and because it is inefficient and self-defeating to teach Es a complete coding system (memory factors might predominate), the EST test, Decoding, employs a simple and ambiguous code. Simplicity and ambiguity were introduced into the coding system for Decoding by employing a code for letters which does not map one-to-one onto the alphabet. The ambiguity of the code allows for words to be judged according to their ease of encoding or decoding. The change from an unambiguous code to an ambiguous one also changes the type of evaluation test involved. Whereas an unambiguous code and experience call for sensitivity, an ambiguous code calls for estimation; the code provides incomplete information, and E must estimate the complete information.

In the test Decoding, E is presented with two words and is asked to choose which one, if coded, would be easier to decode unambiguously. E is also given the opportunity to judge both words as equal in difficulty of decoding.

Jumbled Words is the only test designed for EST that is in the sensitivity category. E is given a stimulus word containing between five and seven letters and is to judge whether or not each of five alternative words is an accurate anagrammatic derivation from the given word. Jumbled Words is, therefore, similar in stimulus material to the test Derivations, which also uses anagram-type stimuli. Pretesting indicated that Jumbled Words correlated higher with EST tests than with Derivations. It was thought that this correlational pattern was due to the fact that Jumbled Words employed all the letters of each word, so that the whole word had to be evaluated for errors in transformation. This criterion is not present in Derivations, wherein only the same letters as those in the given word were the criterion, and fewer could be used.

The third test designed for EST, Typing Errors, is similar to Decoding in the task involved and the stimuli used. E is given an incorrectly typed word and is to choose from among alternatives the word that the incorrectly typed word would most likely be. The judgments are made on the basis of common typing errors due to the arrangement of the typewriter keyboard. A keyboard diagram is printed on each test page for E's reference.

The estimation process involved in responding to Typing Errors is probably not dependent upon EST ability alone, however. It would seem that some figural ability would be involved in this test due to the spatial nature of the keyboard arrangement. Further, it might be expected that typing experience might enter into proficiency at the required task. To determine to some degree the relative contributions of these two extraneous influences, it was decided to include at least one figural factor in the factor analysis and also to correlate test scores with a measure of typing experience.

ESI Tests

Tests designed to measure the factor ESI, evaluation of symbolic implications, employed all types of symbolic stimuli in the list. For the evaluation process, implications are defined as the expectancies or probable relative values of the presented symbols (estimation), or possible symbolic interpretations of a unit or system (sensitivity to symbolic problems). Previous investigations of symbolic-implications factors have employed the following stimuli in the tests: crossword puzzles and disordered symbols (CSI); simple numerical problems (MSI); words and equations (DSI); and simple equations (NSI) (Guilford and Hoepfner, 1963). The ESI tests developed for this study were based upon the stimuli used successfully previously and upon novel stimuli that appeared appropriate for evaluative processing.

The test Abbreviations presented E with a shortened spelling of a common word, E to choose one of the three alternative words that the abbreviated word most likely implies. The meanings of the words are irrelevant to choosing an alternative, and the spelling of the alternatives is correct. The only task for E is to choose the most expected value for the abbreviation, a task of estimating. No observance of shorthand principles was exercised in test construction; the abbreviations were short and relatively unambiguous. Usually, but not always, this implied dropping vowels and unsounded consonants from the keyed alternative. E was warned, however, that sounding-out the abbreviation would not necessarily aid him in his choice.

Abbreviations had been employed in a previous study at the University of Southern California (Petersen, et al., 1963), but was not factor analyzed. Its highest correlations in the two analyses reported were with the test Correlate Completion II, a test of factor NSR. Its reliability was reported as .47. Abbreviations was item-analyzed and lengthened for this investigation in order to improve its reliability.

Letter Problems is similar in format to Form Reasoning, a test of NSI. In Form Reasoning, E makes appropriate substitutions of symbols and solves a simple equation. The evaluation form uses letters as the stimuli and asks E not to solve the equation, but to judge the difficulty or possibility of solving it, on the basis of provided rules. It was hypothesized that E would have to make his judgments based on foresight. E's judgments were of the three-category type; problems were easy to solve (straightforward), difficult to solve (involving manipulations), or impossible to solve due to inadequacies of the table of substitutions.

The third ESI test is named S Test. The "S" in the name means that the test is symbolic. In this test E is given a stimulus about which he is to find a problem to solve. The solution indicates the nature of the problem to which E was sensitive.

The S Test is an adaptation of the F Test (Wilson, et al., 1954), designed as a test of sensitivity to problems, which had its highest (but insignificant) loading on the originality (DMT) factor instead. Although its behavior was not according to hypothesis, it was decided to revise the test to lead E into symbolic responses in all items; some were previously verbal. Then the test might measure E's sensitivity to symbolic implications of unstructured problems. It should be noted, however, that this test is not congruent with the conception that evaluation is "sensitivity to error," for no error is judged. It is a test of E's sensitivity to implications (as, indeed, it turned out) rather than a sensitivity to errors in implications.

PROCEDURES

The Sample

The sample utilized in this study consisted of the entire senior-class student population of the Claremont High School of Claremont, California.[5] Although 131 boys and 180 girls participated in the testing, the sample was later reduced to 86 boys and 139 girls, for whom complete test data for all experimental factor tests were available. The only criterion for exclusion of Es from the sample was incomplete data on these measures.

Age and IQ information was available for 219 and 199 students, respectively. Generalizing from such demographic data available for most of the sample utilized, the estimated mean age was 17.4 years. The estimated mean IQ, computed from combinations of scores obtained from the several IQ measures, which were variously administered between the eighth and eleventh grades, was 110.4. Although IQ's ranged from 80 to 151, no students were deleted from the sample on the basis of extreme indices of general intelligence. The test responses of 14 students with IQ's below 95, who might not have understood the test instructions, were carefully examined to determine whether or not lack of understanding should disqualify them from the sample. None of the students, however, demonstrated consistently poor performance on all tests.

Administration of the Test Battery

The total sample of Es was tested in the mornings and afternoons of Tuesday, November 26, and Wednesday, November 27, 1963. Each testing session required approximately two hours. The tests were arranged into 16-page booklets of about six tests each. The tests of each factor were so arranged that order effects and fatigue effects would be approximately equal for all factors expected to be demonstrated. Each booklet required approximately 50 minutes for administration, so that within each two-hour testing session, there was sufficient time for distributing, administering, and collecting two booklets. Short recesses were given between booklets.

In both of two separate test administrations, the tests were administered at the same times and in the same order. The two examiners used administration manuals which contained introductory and orientation statements, the printed test instructions, and suggested answers to questions expected to be raised by the Es. The test instructions were read aloud by the administrators while the Es maintained pace by reading the instructions silently.

The testing conditions under which the battery was administered were almost ideal, with one major exception. The days of the test administration, November 26 and 27, unfortunately, were only four and five days after the tragic death of President John Fitzgerald Kennedy. Both administrators and school personnel were aware that after the day of national mourning, the preceding Monday, the students were still disturbed and restless. The effects of the national tragedy upon the results of this study are unknown.

[5]For the splendid cooperation in testing in the Claremont High School, we are indebted to Mr. R. H. Oyler, Mr. W. D. Wiley, and Miss Catherine Flannelly.

Scoring

Scoring criteria for the marker tests were developed from the scoring guides employed in previous studies with one exception. Scoring criteria for the newly developed experimental tests were based upon preliminary results of pretestings with university students in undergraduate psychology courses. The S Test was scored subjectively by a trained scorer to determine whether or not each response indicated E's sensitivity to a symbolic implication. Every test was rescored by a second scorer.

The Appendix describes in more detail the scoring procedures for each test. It will be noted that most of the tests, in which alternative responses are made or accepted, have a scoring formula applied to the number of right and wrong responses. Henceforth, such formula scores will be referred to as raw scores, as all analyses were performed upon those scores.

STATISTICAL ANALYSIS AND SOME RESULTS

Statistical Treatment of the Score Data

After scoring, the part and total scores for each test were punched onto IBM cards. The two rating variables, one containing 47 separate ratings, the other containing 8 separate ratings, were also punched onto IBM cards. All school measures available on the Es were punched onto another deck of cards.

Frequency distributions were obtained for all part and total scores to determine whether or not the variables would meet the requirements of the Pearson-r coefficient. A normalizing transformation was applied to those variables that were moderately skewed or exceedingly platykurtic. Extremely skewed or truncated variables were dichotomized near their medians. Descriptions of the frequency distributions and transformations for all the variables are listed in Table 2.

If the four transformed motivation measures actually do measure motivational differences over a large range, one might expect the optimum principle of motivation to be working (Vinacke, 1960). The optimum principle states that performance is best when motivation is at an intermediate or optimal level. As motivation decreases, performance decreases due to lack of interest and attention, and as motivation increases, performance decreases due to the interference of anxiety. This principle suggests that the relationships between the motivation measures and the remaining measures of intellectual performance may be curvilinear. To check this possibility before intercorrelations were computed, scatter plots of several intelligence-variable scores with the motivation-variable scores were inspected. No curvilinearity was apparent from visual inspection, so it was assumed that the optimum principle was not at work to affect the data, probably because no E's motivation was so high as to cause emotional interference.

Table 2

Distributions, Means, Standard Deviations, and Reliabilities of Scores

Test	Form of Distribution[a]	Mean	Standard Deviation	Reliability[d]
1. Abbreviations	0	11.67	4.20	.26[e]
2. Best Letter Set	-	13.47	5.41	.56
3. Best Number Class	--	22.67[b]	6.74	.87
4. Best Number Pairs	0	17.32	5.99	.73
5. Camouflaged Words	0	8.33	2.89	.74[e]
6. Circle Reasoning	-	6.84	2.72	.67[e]
7. Correct Letter Orders	+	16.04	8.17	.58
8. Correct Number Series	0	19.66	9.67	.74
9. Correct Spelling	0	36.98	11.49	.75
10. Decoding	0	16.20	6.39	.74
11. Derivations	0	99.96	20.08	.81
12. Disemvowelled Words	0	11.53	4.21	.79[e]
13. Familiar Letter Combinations	0	14.23	6.59	.43[e]
14. Form Reasoning	--	17.49[b]	5.26	.96[e]
15. Identical Forms	0	38.12	7.30	.63[f]
16. Jumbled Words	--	36.83[b]	10.86	.75
17. Letter Problems	-	14.95[c]	8.88	.88
18. Letter Triangle	+	5.68	2.85	.55[e]
19. Letter "U"	0	53.63	11.28	.84
20. Marking Speed	+	94.24[c]	18.12	.44[f]
21. Number Classification	-	11.41[c]	3.35	.72[e]
22. Number Grouping	-	14.10	4.42	.79
23. Number-Group Naming	--	10.17[b]	2.12	.76[e]
24. Numerical Operations	+	22.23	8.62	.78
25. Operations Sequence	0	12.27	5.29	.80
26. Perceptual Speed	0	48.01	9.21	.65[f]
27. Related Words I	-	14.00	5.41	.55
28. S Test	+	8.55	3.42	.69[e]
29. Seeing Trends II	0	8.16	3.25	.80[e]
30. Series Relations	0	9.93	7.22	.74
31. Sign Changes	-	17.49	5.05	.57
32. Sign Changes II	--	17.31[b]	3.64	.82
33. Similar Pairs	-	20.14[c]	7.40	.74
34. Sound Grouping	+	11.92	6.59	.74
35. Symbol Grouping	0	11.22	4.84	.84[g]
36. Symbol Identities	0	72.12	14.21	.90
37. Symbol Manipulation	--	21.40[b]	8.27	.74
38. Symbol Reasoning	+	17.95[c]	10.29	.78
39. Typing Errors	+	9.44[c]	4.99	.49
40. Varied Symbols	0	10.92	4.50	.67
41. Way-Out Numbers	--	23.48[b]	6.55	.76
42. Word Changes	--	11.07[b]	4.47	.87[g]
43. Word Choice	0	13.05	5.81	.62
44. Word Combinations	+	10.05[c]	6.42	.72
45. Word Patterns	0	70.39	9.14	.75
46. Word Relations	-	10.76	4.64	.78
47. Word Transformation	-	26.69	7.91	.83[g]
48. Rating - Test Liking	0	100.01[c]	20.71	.90
49. Rating - Booklet Effort	-	200.06[c]	65.28	.95
50. Sex	+	.38	.49	--
51. Prediction-Achievement Discrepancy	0	199.42	59.13	.43[f]
52. ITED General Vocabulary	-	20.35	5.45	.81[f]
53. PSAT Verbal	0	47.88	11.62	.95[f]
54. SCAT Verbal	0	304.82	14.33	.85[f]
55. Mathematics Experience	+	3.63	2.22	--
56. Typing Experience	+	1.26	.81	--
57. Shorthand Experience	++	.24	.54	--
58. Test Administrator	+	.33	.47	--

[a] The skewness of score distributions is described as follows: ++, strong positive skewness; +, slight positive skewness; 0, normal distribution; -, slight negative skewness; and --, strong negative skewness.

[b] Total scores dichotomized at the medians for intercorrelations.

[c] Total scores C-scaled for intercorrelations.

[d] All estimates of reliability are Spearman-Brown corrections of correlation between parts unless noted.

[e] Kuder-Richardson estimate of reliability.

[f] Communality entered as reliability estimate.

[g] Reliability estimated through formula 21.21, in Gulliksen(1950).

After it was ascertained that all the part-score data and total-score data, raw or scaled, met the requirements of the Pearson r or its approximations, reliability estimates were obtained from the raw scores of the tests. For all tests with two or more parts, Spearman-Brown reliability estimates were computed. Kuder-Richardson estimates of reliability were computed for all one-part tests that showed no evidence of speeding. Reliability estimates for one-part tests, wherein each item had a large possible range of scores, were computed by a formula suggested by Gulliksen (1950, p. 378). Reliabilities of one-part speeded tests and of school measures could not be estimated. Their communalities were expected to approximate the necessary estimates of reliability.

In all cases, the new estimates of test reliabilities were very similar to those obtained from previous studies or from pretesting results where tests had been used before. This is especially interesting in the cases of the marker tests, which had been generally shortened. The reliabilities reported in Table 2 suggest that shortening a test on empirical bases does not, upon readministration, severely reduce its reliability.

Item analyses of two experimental tests that retained low reliability estimates from pretesting (Abbreviations and Familiar Letter Combinations) did not improve the estimates, so the tests were included in the analysis without item deletions. All of the reliability estimates reported in Table 2, therefore, are based upon the tests exactly as they were printed and administered, with no after-the-fact item deletions or item weightings. In general, the reliability estimates for all tests were considered adequate for the factor-analysis procedures to follow.

The means and standard deviations of the variables are also listed in Table 2. In all but three cases, these descriptive statistics are based upon raw scores, before any transformations were applied. Variables 48, 49, and 51 were automatically standardized in their computations so that their means and standard deviations have been arbitrarily set. Few comparisons between the descriptive statistics computed from this analysis and those from other analyses can be made, either because the tests were shortened and otherwise altered, or else because they had never before been administered to a large sample.

Intercorrelations

The score matrix of 58 raw, scaled, and dichotomized variables was punched onto IBM cards. Because the score matrix was incomplete and the data were differentially scaled, resulting in different kinds of correlation coefficients, the correlation matrix was obtained from the WDCORR program at the Western Data Processing Center. This program computes correlation coefficients between variables based upon the total number of individuals for whom scores are available. Most of the correlation coefficients in Table 3, therefore, are based upon the whole sample of 225 Es, but some are based upon the four variables for which not all Es' scores were available. These variables and the number of scores available are: variable 51, 165 scores; variable 52, 187 scores; variable 53, 109 scores; and variable 54, 107 scores. The attenuated samples for these variables taken independently, of course, result in further attenuation of sample size for the coefficients among them. The coefficient between variables 52 and 54 was computed from a common sample of only 66 Es. This sample size was the smallest from which any coefficients were computed and was considerably smaller than the next smallest sample of 87 for the correlation between variables 52 and 53.

Such variations in sample size upon which correlation coefficients are based introduces additional possibilities of error into the correlation matrix due to the necessary generalization that each coefficient in the matrix equally estimates the actual intercorrelation between the variables. It was decided, however, that a biased estimate would be better than none at all, and since the Es in the reduced samples appeared to have been selected on irrelevant variables, i. e., their attendance at the school when the tests were administered, the bias was thought to be small.

An additional consideration of the WDCORR program is that it computes product-moment correlation coefficients upon any input data. This means that some coefficients are point-biserial r's and some are phi coefficients. Standard corrections were applied to each kind of coefficient to improve it as an estimate of the Pearson r. Thus, the coefficients reported in Table 3 are all Pearson-r coefficients or estimates of the Pearson r.

The correlation matrix presented in Table 3 is typical of those obtained from intellectual measures, with some exceptions. The introduction of motivation measures (variables 20, 48, 49, and 51) and experience measures (variables 55, 56, 57, and 58) resulted in a large number of negative coefficients not generally found among measures of intelligence. It will be seen that when only the intellectual measures are considered, however, there is still a considerable number of negative coefficients, some as large as -.13. These coefficients are pointed out, since they will influence the positive-manifold requirements of the factor matrix, discussed in the next section.

Before discussing the factor analysis, it is appropriate to discuss the five variables not to be factor analyzed. These variables, numbers 50, 55, 56, 57, and 58, were included in the correlational analysis to determine their relationships to particular intellectual test scores. Examining the correlations with variable 50, sex, it can be seen that 15 variables correlate with sex at the .05 level of significance, and that eight variables correlate at the .01 level. Of these eight variables, two are such that boys tend to have higher scores. They are variables 30, Series Relations, and 55, Mathematics Experience. In the remaining six variables, girls tend to perform better. It is interesting that most of these variables are tests like clerical- and perceptual-speed tests. Girls excel on variables 9, Correct Spelling; 11, Derivations; 16, Jumbled Words; 26, Perceptual Speed; 36, Symbol Identities; 56, Typing Experience; and 57, Shorthand Experience. All of the significant correlations with sex are reasonable.

Table 3

Correlation Matrix[a] of 58 Variables (N = 225)[b]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Abbreviations		32	33	40	14	36	35	38	31	36	27	17	24	14	26	30	36	27	19	-02	38	36	33	11
2. Best Letter Set	32		50	44	17	21	40	46	26	42	28	17	20	33	27	28	35	37	23	05	45	45	48	27
3. Best Number Class	33	50		65	16	28	53	66	45	57	33	21	23	25	30	38	34	40	31	10	62	62	57	39
4. Best Number Pairs	40	44	65		20	28	54	66	41	51	21	16	22	28	28	50	46	41	18	-03	59	64	69	19
5. Camouflaged Words	14	17	16	20		10	20	22	24	19	13	30	07	-10	15	18	12	12	14	13	11	23	26	10
6. Circle Reasoning	36	21	28	28	10		25	33	28	28	19	21	19	18	16	38	20	31	09	02	22	33	33	-08
7. Correct Letter Orders	35	40	53	54	20	25		59	36	46	28	19	13	23	19	46	40	34	17	-02	42	51	53	24
8. Correct Number Series	38	46	66	66	22	33	59		41	49	32	22	23	28	36	59	43	47	29	08	54	62	58	33
9. Correct Spelling	31	26	45	41	24	28	36	41		35	26	44	21	33	19	45	25	13	18	03	40	39	39	26
10. Decoding	36	42	57	51	19	28	46	49	35		26	17	17	15	25	41	36	27	18	02	45	53	53	22
11. Derivations	27	28	33	21	13	19	28	32	26	26		21	22	32	34	49	25	25	33	20	33	34	28	33
12. Disemvowelled Words	17	17	21	16	30	21	19	22	44	17	21		16	14	19	35	13	19	26	05	18	18	23	17
13. Familiar Letter Combinations	24	20	23	22	07	19	13	23	21	17	22	16		25	18	28	21	15	10	11	21	20	28	11
14. Form Reasoning	14	33	25	28	-10	18	23	28	33	15	32	14	25		23	31	29	24	19	13	24	29	24	29
15. Identical Forms	26	27	30	28	15	16	19	36	19	25	34	19	18	23		26	25	20	33	21	28	34	30	19
16. Jumbled Words	30	28	38	50	18	38	46	59	45	41	49	35	28	31	26		41	44	33	19	39	50	39	41
17. Letter Problems	36	35	34	46	12	20	40	43	25	36	25	13	21	29	25	41		36	34	03	26	42	30	10
18. Letter Triangle	27	37	40	41	12	31	34	47	13	27	25	19	15	24	20	44	36		20	-04	36	32	44	19
19. Letter U	19	23	31	18	14	09	17	29	18	18	33	26	10	19	33	33	34	20		38	26	28	26	36
20. Marking Speed	-02	05	10	-03	13	02	-02	08	03	02	20	05	11	13	21	19	03	-04	38		01	11	04	33
21. Number Classification	38	45	62	59	11	22	42	54	40	45	33	18	21	24	28	39	26	36	26	01		54	53	30
22. Number Grouping	36	45	62	64	23	33	51	62	39	53	34	18	20	29	34	50	42	32	28	11	54		58	20
23. Number-Group Naming	33	48	57	69	26	33	53	58	39	53	28	23	28	24	30	39	30	44	26	04	53	58		18
24. Numerical Operations	11	27	39	19	10	-08	24	33	26	22	33	17	11	29	19	41	10	19	36	33	30	20	18	
25. Operations Sequence	28	51	59	52	19	28	41	58	37	45	33	29	18	34	31	46	38	44	35	09	53	50	55	36
26. Perceptual Speed	21	27	33	26	10	15	20	34	14	22	43	08	20	38	56	36	20	22	21	10	28	28	29	15
27. Related Words I	28	43	45	45	09	25	44	43	35	39	41	19	18	34	23	46	36	38	16	-06	44	43	39	19
28. S Test	-02	13	06	12	08	07	12	19	07	12	07	09	-09	00	17	14	15	06	09	01	17	16	14	08
29. Seeing Trends II	30	33	35	49	24	25	57	48	32	39	19	13	20	29	22	45	40	38	08	-05	34	42	40	13
30. Series Relations	27	36	50	47	15	27	50	50	35	44	20	26	10	20	21	48	32	37	21	07	38	53	45	30
31. Sign Changes	21	37	50	37	18	12	28	45	26	35	41	17	13	30	36	41	21	28	47	34	38	41	42	56
32. Sign Changes II	26	30	53	49	10	19	36	45	28	43	19	17	16	14	26	34	32	38	18	00	40	41	60	27
33. Similar Pairs	36	42	43	50	08	30	42	48	34	33	30	22	22	33	21	53	41	47	19	-11	36	41	38	11
34. Sound Grouping	33	41	48	49	29	29	43	50	51	42	23	42	18	23	23	35	35	28	13	-01	38	42	45	11
35. Symbol Grouping	29	38	54	44	06	30	41	56	22	40	29	13	18	27	26	39	38	36	26	01	45	48	45	21
36. Symbol Identities	24	24	30	34	09	08	30	41	42	30	47	20	09	37	50	49	32	21	51	27	33	40	31	41
37. Symbol Manipulation	25	38	68	48	19	33	40	50	26	38	31	15	16	30	19	38	29	45	14	-11	46	53	53	06
38. Symbol Reasoning	33	42	62	49	13	36	56	60	33	44	30	15	18	22	28	41	39	42	18	-02	43	51	53	17
39. Typing Errors	13	29	29	37	14	16	28	28	20	30	20	15	11	32	15	29	23	26	17	01	33	34	28	13
40. Varied Symbols	03	11	-01	09	12	07	12	17	09	16	26	27	02	09	19	19	07	11	12	-02	08	13	11	13
41. Way-Out Numbers	25	24	59	46	08	14	42	53	23	29	25	13	10	35	30	36	26	33	20	11	35	41	49	41
42. Word Changes	48	39	49	47	19	34	49	56	29	44	29	34	20	30	33	58	34	53	32	05	42	51	66	29
43. Word Choice	27	34	39	41	02	34	36	46	22	41	32	12	14	29	17	41	40	30	15	00	37	47	44	17
44. Word Combinations	29	35	35	36	23	30	48	45	40	35	23	37	23	27	22	45	31	28	15	02	29	39	44	20
45. Word Patterns	25	20	33	30	25	28	30	36	26	26	29	27	18	23	27	38	31	27	24	08	31	40	43	15
46. Word Relations	32	37	48	54	20	29	49	57	35	40	39	25	27	27	39	59	40	43	14	03	43	47	49	25
47. Word Transformation	31	36	54	40	33	33	41	42	49	39	33	46	23	11	24	44	23	32	19	09	42	44	40	27
48. Rating Test Liking	15	20	23	21	19	14	19	18	21	15	15	24	09	-01	10	20	19	24	14	06	22	28	21	05
49. Rating Booklet Effort	05	-09	-08	-02	03	02	-05	-06	06	-06	11	03	-05	03	-06	-06	-03	08	01	-03	-02	-01	-13	-01
50. Sex	-11	-08	09	-06	-09	-08	-02	00	-18	01	-31	-09	-11	-11	-10	-19	-15	-01	-03	-01	-04	-08	-02	10
51. Prediction-Achievement Discrepancy	15	11	10	18	-18	02	14	09	05	13	-02	-10	-07	14	01	13	08	03	11	-05	08	17	10	10
52. ITED General Vocabulary	36	41	54	56	11	30	65	56	49	43	16	20	16	20	25	36	45	34	11	-10	36	48	52	06
53. PSAT Verbal	30	25	39	40	39	23	51	52	47	46	14	20	13	15	27	56	28	25	11	-03	26	42	31	04
54. SCAT Verbal	27	38	30	41	17	11	47	47	44	29	08	30	07	32	15	34	21	23	-05	-19	31	37	40	06
55. Mathematics Experience	25	31	52	50	11	27	48	56	28	41	04	13	05	16	18	33	27	28	15	04	39	51	57	18
56. Typing Experience	-02	-20	-28	-13	06	00	-14	-09	-01	-19	12	10	-04	-10	00	06	-13	-04	04	15	-09	-15	-18	07
57. Shorthand Experience	-08	-03	-16	-09	11	-07	-12	-15	00	-10	14	15	09	-10	06	00	-09	-05	01	13	01	-05	-06	00
58. Test Administrator	-07	-07	04	12	-24	11	16	09	01	13	09	-09	-02	09	-06	26	04	06	-04	-03	02	13	-04	-02

[a]Decimal points omitted.

[b]A sample of 225 students determined the coefficients for variables 1 through 50 and variables 55 through 58. Coefficients for variables 51, 52, 53, and 54 were determined respectively from subsamples of 165, 182, 109, and 107 as school records were incomplete.

Variable 55, Mathematics Experience, was included in the correlational analysis to determine which, if any, of the aptitude-test scores could possibly depend upon the amount of mathematical experience E had had. Fifty of the variables correlate at the .05 level of significance with the index of mathematics experience. This large number of significant correlation coefficients could be due to two things: (1) experience in mathematics courses trains individuals at proficiency in tasks measured by symbolic tests, or (2) individuals who get more mathematics training are relatively high in abilities measured by the tests. It is probable that both explanations underlie the coefficients, but the second would have greater practical significance. If the second explanation is correct, the tests correlating with mathematics experience show promise as selectors of mathematics students.

To determine whether or not scores on test 39, Typing Errors, are related to experience in typing, the number of semesters of formal high-school typing courses taken was entered into the correlational analysis. The two variables, numbers 39 and 56, correlate .03, signifying that there is little or no relationship between them. It is interesting to note that Typing Experience is correlated positively and significantly only with Disemvowelled Words (a CSU test), in which word spelling and recognition are important, with Marking Speed, a measure of finger speed, and with Shorthand Experience. It is significantly negatively correlated with most of the numerical tests and with all three verbal-comprehension tests.

Variable 57, Shorthand Experience, was analyzed to determine its relationship with variable 1, Abbreviations. The two variables under consideration correlate negatively, but not significantly. Shorthand Experience correlates positively with word and speed tests and with Typing Experience, but correlates negatively with many numerical tests and all three verbal-comprehension tests.

Table 3 (Continued)

25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
28	21	28	-02	30	27	21	26	36	33	29	24	25	33	13	03	25	48	27	29	25	32	31	15	05	-11	15	36	30	27	25	-02	-08	-07
51	27	43	13	33	36	37	30	42	41	38	24	38	42	29	11	24	39	34	35	20	37	36	20	-09	-08	11	41	25	38	31	-20	-03	-07
59	33	45	06	35	50	50	53	43	48	54	30	68	62	29	-01	59	49	39	35	33	48	54	23	-08	09	10	54	39	30	52	-28	-16	04
52	26	45	12	49	47	37	49	50	49	44	34	48	49	37	09	46	47	41	36	30	54	40	21	-02	-06	18	56	40	41	50	-13	-09	12
19	10	09	08	24	15	18	10	08	29	06	09	19	13	14	12	08	19	02	23	25	20	33	19	03	-09	-18	11	39	17	11	06	11	-24
28	15	25	07	25	27	12	19	30	29	30	08	33	36	16	07	14	34	34	30	28	29	33	14	02	-08	02	30	23	11	27	00	-07	11
41	20	44	12	57	50	28	36	42	43	41	30	40	56	28	12	48	49	36	42	30	49	41	19	-05	-02	14	65	51	47	48	-14	-12	16
58	34	43	19	48	50	45	45	48	50	56	41	50	60	28	17	53	56	46	45	36	57	42	18	-06	00	09	56	52	47	56	-09	-15	09
37	14	35	07	32	35	26	28	34	51	22	42	26	33	20	09	23	29	22	40	26	35	49	21	06	-18	05	49	47	44	28	-01	00	01
45	22	39	12	39	44	35	43	33	42	40	30	38	44	30	16	29	44	41	35	26	40	39	15	-06	01	13	43	46	29	41	-19	-10	13
33	43	41	07	19	20	41	19	30	23	29	47	31	30	20	26	25	29	32	23	29	39	33	15	11	-31	-02	16	14	08	04	12	14	09
29	08	19	09	13	26	17	17	22	42	13	20	15	15	15	27	13	34	12	37	27	25	46	24	03	-09	-10	20	20	30	13	10	15	-09
18	20	18	-09	20	10	13	16	22	18	18	09	16	18	11	02	10	20	14	23	18	27	23	09	-05	-11	-07	16	13	07	05	-04	09	-02
34	38	34	00	29	20	30	14	33	23	27	37	30	22	32	09	35	30	29	27	23	27	11	-01	03	-11	14	20	15	32	16	-10	-10	09
31	56	23	17	22	21	36	26	21	23	26	50	19	28	15	19	30	33	17	22	27	39	24	10	-06	-10	01	25	27	15	18	00	06	-06
46	36	46	14	45	48	41	34	53	35	39	49	38	41	29	19	36	58	41	45	38	59	44	20	-06	-19	13	36	56	34	33	06	00	26
38	20	36	15	40	32	21	32	41	35	38	32	29	39	23	07	26	34	40	31	31	40	23	19	-03	-15	08	45	28	21	27	-13	-09	04
44	22	38	06	38	37	28	38	47	28	36	21	45	42	26	11	33	53	30	28	27	43	32	24	08	-01	03	34	25	23	28	-04	-05	06
35	21	16	09	08	21	47	18	19	13	26	51	14	18	17	12	20	32	15	15	24	14	19	14	01	-03	11	11	11	-05	15	04	01	-04
09	10	-06	01	-05	07	34	00	-11	-01	01	27	-11	-02	01	-02	14	05	00	02	08	03	09	06	-03	-01	-05	-10	-03	-19	04	15	13	-03
53	28	44	17	34	38	38	40	36	38	45	33	46	43	33	08	35	42	37	29	31	43	42	22	-02	-04	08	36	26	31	39	-09	01	02
50	28	43	16	42	53	41	41	41	42	48	40	53	51	34	13	41	51	47	39	40	47	44	28	-01	-08	17	48	42	37	51	-15	-05	13
55	29	39	14	40	45	42	60	38	45	45	31	53	53	28	11	49	66	44	44	43	49	40	21	-13	-02	10	52	31	40	57	-18	-06	-04
35	15	19	08	13	30	56	27	11	11	21	41	06	17	13	13	41	29	17	20	15	25	27	05	-01	10	10	06	04	06	18	07	00	-02
	26	45	16	30	55	50	49	42	47	45	37	50	55	40	21	38	47	39	33	41	49	37	29	-06	13	07	37	28	37	51	-18	-12	-01
26		31	16	23	15	29	19	27	19	37	32	19	28	15	12	19	23	23	18	24	34	19	03	-03	-18	08	22	26	24	16	-02	00	06
45	31		04	35	36	29	35	50	40	48	30	49	41	32	14	33	44	40	25	19	59	27	22	02	-14	-02	43	26	33	22	-12	-02	12
16	16	04		12	07	06	16	04	08	18	17	00	11	10	22	10	04	11	08	16	09	09	09	-13	-03	-09	16	-09	01	30	08	02	-15
30	23	35	12		41	22	28	44	38	31	22	44	47	22	14	35	49	31	36	32	52	36	22	02	-09	07	54	55	46	35	-14	-15	13
55	15	36	07	41		38	36	41	39	47	24	45	51	31	11	48	54	32	35	27	48	38	27	00	18	10	42	32	44	46	-17	-14	11
50	29	29	06	22	38		33	24	17	33	51	16	29	21	16	44	40	20	17	27	34	26	12	-08	07	06	17	12	02	27	-01	05	-05
49	19	35	16	28	36	33		28	41	42	25	39	40	23	18	39	43	33	26	26	41	33	14	-04	06	-02	38	19	30	45	-19	-20	-03
42	27	50	04	44	41	24	28		40	47	34	51	48	35	09	34	49	41	37	26	55	33	21	01	-15	07	50	33	41	31	-06	-03	10
47	19	40	08	38	39	17	41	40		34	26	40	44	25	15	26	43	30	44	33	47	47	22	-09	-12	-12	60	52	56	37	-10	-08	02
45	37	48	18	31	47	33	42	47	34		27	43	44	26	09	36	45	39	32	34	43	27	16	-04	-02	-05	44	23	19	43	-08	-09	06
37	32	30	17	22	24	51	25	34	26	27		14	31	19	18	33	38	26	20	26	35	24	09	00	-25	13	35	26	11	23	10	08	08
50	19	49	00	44	45	16	39	51	40	43	14		56	30	03	30	48	44	26	39	45	36	18	01	-03	01	43	44	50	33	-25	-10	05
55	28	41	11	47	51	29	40	48	44	44	31	56		32	11	49	40	38	36	31	54	33	20	-03	05	08	57	52	42	53	-22	-11	16
40	15	32	10	22	31	21	23	35	25	26	19	30	32		11	20	23	19	23	17	31	26	08	-01	-04	07	27	24	19	26	03	04	00
21	12	14	22	14	11	16	18	09	15	09	18	03	11	11		03	20	15	13	21	26	20	05	-06	-03	-13	10	00	05	16	04	11	01
38	19	33	10	35	48	44	39	34	26	36	33	30	49	20	03		49	25	25	14	40	24	14	-04	11	04	35	30	30	33	-14	-09	18
47	23	44	04	49	54	40	43	49	43	45	38	48	40	23	20	49		38	40	34	58	49	19	01	-14	04	43	52	34	40	-05	-09	01
39	23	40	11	31	32	20	33	41	30	39	26	44	38	19	15	25	38		32	36	44	27	16	-03	-10	02	37	24	35	31	-10	-09	10
33	18	25	08	36	35	17	26	37	44	32	20	26	36	23	13	25	40	32		31	43	47	22	06	-14	00	39	40	32	32	-08	00	03
41	24	19	16	32	27	27	26	26	33	34	26	39	31	17	21	14	34	36	31		34	41	30	-01	03	-02	30	27	12	35	00	-02	-09
49	34	59	09	52	48	34	41	55	47	43	35	45	54	31	26	40	58	44	43	34		40	25	02	-06	-05	46	32	31	38	-13	-11	16
37	19	27	09	36	38	26	33	33	47	27	24	36	33	26	20	24	49	27	47	41	40		25	05	-05	00	43	44	45	31	-02	01	05
29	03	22	09	22	27	12	14	21	22	16	09	18	20	08	05	14	19	16	22	30	25	25		35	-04	-12	21	11	05	22	-04	00	05
-06	-03	02	-13	02	00	-08	-04	01	-09	-04	00	01	-03	-01	-06	-04	01	-03	06	-01	02	05	35		-09	-03	-02	04	-12	-07	09	-01	14
13	-18	-14	-03	-09	18	07	06	-15	-12	-02	-25	-03	05	-04	-03	11	-14	-10	-14	03	-06	-05	-04	-09		-03	-03	-06	06	30	-35	-35	-01
07	08	-02	-09	07	10	06	-02	07	-12	-05	13	01	08	07	-13	04	04	02	00	-02	-05	00	-12	-03	-03		11	23	18	17	-09	-06	12
37	22	43	16	54	42	17	38	50	60	44	35	43	57	27	10	35	43	37	39	30	46	43	21	-02	-03	11		71	63	57	-30	-22	12
28	26	26	-09	55	32	12	19	33	52	23	26	44	52	24	00	30	52	24	40	27	32	44	11	04	-06	23	71		66	50	-18	-19	06
37	24	33	01	46	44	02	30	41	56	19	11	50	42	19	05	30	34	35	32	12	31	45	05	-12	06	18	63	66		37	-39	-33	-02
51	16	22	30	35	46	27	45	31	37	43	23	33	53	26	16	33	40	31	32	35	38	31	22	-07	30	17	57	50	37		-14	-24	-01
-18	-02	-12	08	-14	-17	-01	-19	-06	-10	-08	10	-25	-22	03	04	-14	-05	-10	-08	00	-13	-02	-04	09	-35	-09	-30	-18	-39	-14		48	-09
-12	00	-02	02	-15	-14	05	-20	-03	-08	-09	08	-10	-11	04	11	-09	-09	-09	00	-02	-11	01	00	-01	-35	-06	-22	-19	-33	-24	48		-05
-01	06	12	-15	13	11	-05	-03	10	02	06	08	05	16	00	01	18	01	10	03	-09	16	05	05	14	-01	12	12	06	-02	-01	-09	-05	

Only three tests correlate at the .01 level of significance with variable 58, Test Administrator. For unexplainable reasons, the students in the smaller group tended to score higher on test 16, Jumbled Words, and 41, Way-Out Numbers. Students in the larger group tended to score higher on test 5, Camouflaged Words. In general, however, the correlations with variable 58 were low.

The five variables just discussed are either not appropriate for factor analysis or would have confused the hypothesized factor structure due to their experiential natures. The high correlations across the board with Mathematics Experience might generate a g-type factor, and, similarly, other experience indexes, correlated with composites of tests, might generate group-factors, thus militating against a clear test of the hypotheses derived from the structure-of-intellect model. Discussion of the five variables ends here.

Factor Analysis

The correlation matrix of all variables except numbers 50, 55, 56, 57, and 58 was submitted to an iterative communality-estimation program for estimates of communalities to be inserted into the principal-diagonal cells of the correlation matrix for factor analysis. Starting with the highest-in-column entries, the iterative program extracted 19 factors (the hypothesized number of factors), computed the new communalities, replaced the diagonal entries with the new communalities, and re-extracted factors. This procedure was iterated eight times until the communality estimates stabilized from one iteration cycle to the next.

The iterated, stabilized communality estimates were put into the diagonal cells of the correlation matrix and the matrix was submitted to the BIMD 17 program at the Western Data Processing Center. BIMD 17 extracts principal-axes factors until the eigenvalues become negative, at which point extractions are stopped.

Table 4

Unrotated Factor Matrix[a]

Var.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	h^2
1.	50	05	-06	-01	-11	08	24	18	10	14	19	00	-06	-08	-10	18	07	-08	-10	-11	52
2.	59	-02	11	-02	03	-03	-13	20	03	-03	-05	-21	-06	21	-11	25	07	-08	-03	00	61
3.	77	-05	21	00	30	07	-02	19	-20	-06	04	16	08	-05	-02	09	-04	11	01	02	87
4.	75	11	19	02	12	-01	08	10	13	-03	12	-02	-12	07	-06	-21	-08	09	00	-04	76
5.	28	05	-42	-08	22	-11	14	-03	-19	-04	-16	-06	-05	19	-14	-06	06	06	06	-06	49
6.	42	11	-07	-22	-15	03	14	14	04	19	-02	11	08	-08	08	03	-15	-15	-10	-01	43
7.	69	18	07	13	07	-02	06	-17	11	-09	01	08	-03	10	07	11	08	-02	-05	-10	64
8.	80	00	08	07	08	-07	07	-07	-01	01	01	05	06	03	05	03	-07	06	-04	01	69
9.	56	08	-35	15	05	09	-20	18	11	-12	17	13	04	-08	02	-11	-05	02	-02	-05	65
10.	64	04	04	00	15	-07	08	08	13	06	-03	18	-12	06	-15	01	03	-13	07	16	60
11.	46	-37	-10	-06	-30	03	-10	01	-09	-07	-06	24	-09	-05	-03	07	14	-03	-06	-03	58
12.	37	-04	-54	-14	04	03	-24	01	04	09	01	-11	03	-14	01	04	-09	02	04	-05	57
13.	30	-06	-07	-07	-14	03	01	20	-12	17	21	03	12	21	-01	-07	01	03	-08	01	32
14.	43	-19	10	18	-34	14	-27	14	-01	02	03	-16	04	09	23	-13	02	-04	-02	03	60
15.	45	-32	-10	12	-16	-34	17	04	-23	-05	10	-16	-08	-10	-01	00	-09	-08	01	00	62
16.	69	-12	-15	11	-21	12	00	-21	10	19	-12	23	02	00	-04	05	-12	11	01	04	77
17.	55	-01	06	-01	-19	-03	17	00	27	-08	-01	-09	21	11	-07	02	08	08	-01	08	54
18.	56	01	12	-21	-10	16	06	-17	-06	09	-06	-20	-04	-05	-06	06	-03	15	-16	08	56
19.	37	-52	-12	08	05	08	15	09	11	-04	-16	-17	18	-13	-10	03	04	08	04	00	61
20.	07	-51	-17	12	12	05	11	06	-09	01	-06	02	20	15	06	-05	02	-09	03	04	43
21.	65	-07	15	-07	13	00	-06	21	02	-08	04	08	-18	-03	-03	00	-04	11	-05	-13	60
22.	75	-01	08	00	09	-01	15	13	07	-06	-12	09	-07	04	09	-09	-03	-12	12	-06	70
23.	75	03	12	-11	22	-12	07	09	00	19	-10	-13	-07	-05	08	-18	10	-03	02	-02	79
24.	36	-52	-03	19	23	20	-21	-16	00	02	06	08	-05	12	06	13	08	21	-02	06	67
25.	71	-14	08	-14	13	-01	-17	05	04	-05	-20	-13	01	-05	-01	-02	-03	-09	-13	07	69
26.	43	-24	04	16	-36	-29	05	13	-29	-03	02	-03	-17	02	07	10	-11	05	06	09	64
27.	62	01	15	-12	-24	07	-20	-04	-04	-12	05	07	-03	01	-23	-02	07	-05	15	-04	63
28.	16	-12	-01	-12	04	-39	-01	-11	22	-12	-06	-02	-07	00	14	05	-10	14	05	-06	35
29.	61	25	-03	09	-09	01	15	-20	-01	-03	02	-08	-07	23	08	-05	09	05	-04	-07	61
30.	65	04	06	-03	18	16	-05	-17	02	01	-10	-05	03	00	09	08	-17	-21	10	-03	63
31.	51	-55	05	09	20	06	00	-03	-05	02	-01	-03	-02	05	-06	-01	03	-08	-04	-05	64
32.	57	-03	15	-16	22	-14	-07	-06	03	08	15	-02	-03	-15	-02	-08	09	08	-01	27	60
33.	64	13	11	-05	-29	15	-06	-06	04	-01	00	-08	07	-01	-12	03	-07	10	03	-10	61
34.	64	25	-26	-02	07	-15	-19	08	00	-05	10	-06	17	-02	-10	-02	06	-06	03	05	68
35.	62	-07	23	-16	-07	-07	04	00	-01	-03	04	03	16	00	03	15	-14	04	21	03	60
36.	51	-45	-10	29	-12	-05	04	-04	16	-16	04	-01	03	-19	-06	-13	07	01	01	-07	70
37.	66	25	20	-18	-03	12	-05	10	-28	02	-27	02	06	-11	04	-08	13	10	06	-07	81
38.	67	13	18	00	02	-12	02	-10	-11	-16	-08	11	13	-07	01	02	-03	-12	-28	00	70
39.	42	-01	07	-02	-04	03	-15	05	07	-07	-20	-08	-07	13	-12	-18	-20	-02	-08	03	38
40.	19	-15	-20	-20	-07	-26	-14	-25	13	06	-08	04	-17	-03	-02	00	06	-07	-06	01	36
41.	56	-13	22	14	18	08	02	-28	-18	-02	18	-01	06	-05	13	-03	-05	-02	-01	-07	61
42.	72	-02	-06	-05	01	15	13	-22	-06	32	06	-16	-05	-13	-06	01	04	-07	11	-05	79
43.	56	04	14	-13	-18	-02	-02	04	14	07	-08	15	03	00	17	00	19	-04	10	08	52
44.	56	10	-26	-06	-04	05	-02	02	08	09	10	-02	02	16	12	05	-10	00	01	10	49
45.	50	-07	-19	-22	-05	-08	17	11	04	05	-13	01	02	-02	23	-04	11	08	-05	-06	49
46.	71	-01	00	-16	-18	-02	-04	-20	-06	01	13	04	-05	11	-08	-11	-02	-05	03	-03	66
47.	61	08	-38	-07	15	05	-06	08	-05	08	00	09	-11	01	04	12	-04	10	-02	-01	62
48.	31	01	-20	-35	06	19	15	-01	01	-34	04	-10	-06	02	10	02	01	-04	04	02	47
49.	-02	02	-16	-18	-12	39	16	-01	-07	-34	10	-02	-20	-08	08	-02	01	-03	00	16	46
51.	09	03	17	42	-01	20	11	17	22	10	-10	-07	-22	-07	02	04	-06	02	01	03	43
52.	70	39	-03	21	-01	-13	07	-03	12	-20	12	-02	14	-07	-01	06	04	02	-01	02	81
53.	61	42	-29	44	-02	04	22	-06	-16	-01	-16	07	02	-02	-10	-04	-01	01	00	13	97
54.	56	51	-08	34	00	-05	-30	-01	-08	02	-08	-13	-05	-07	12	07	09	-02	03	00	85

[a]Decimal points omitted.

This procedure resulted in the extraction of 34 principal-axes factors. The first 19 factors accounted for 93.6 per cent of the total variance of the 34-factor matrix. Inclusion of the 20th factor, which had a range of loadings from -.13 to .27, as a potential residual factor, accounted for 94.4 per cent of the total variance. The 20-factor matrix was accepted as the principal-axes factor matrix to be rotated. With 20 factors, the average input-output communality discrepancy was .029. This factor matrix is presented in Table 4.

Nineteen of the 20 factors of the principal-axes matrix were analytically rotated to quartimax, varimax, and equamax solutions. Since all three rotational criteria tended to make each factor rotated at least a strong singlet, and since one residual factor was expected, the 20th principal-axes factor was withheld from the analytical rotational solutions. The equamax solution appeared more in line with the hypothesized factor structure because of its tendency to equalize approximately the variance on each factor while seeking simple structure.

The 19 equamax-rotated factors, along with the 20th principal-axes factor, were then graphically rotated to orthogonal simple structure and psychological meaningfulness. The rotational criterion of positive manifold was necessarily violated due to the many negative relationships between variables. After 61 rotations, it was clear that most of the hypothesized factors would emerge, but that the criterion of simple structure would not strictly be met without further minor adjustments in the rotations.

The graphic solution was submitted to a program designed to rotate the loadings as closely as possible to a fixed target matrix of loadings (Cliff, 1964). The construction of the target matrix depended upon the intuitively inferred structure of the empirical matrix, simple structure, positive manifold, and the factor hypotheses. Four slight successive adjustments or revisions of the target matrix effected considerable improvement in the empirical rotated matrix on all four criteria. The principal-axes matrix was then rotated to the fifth target matrix so that rounding and graphic errors would

Table 5

Rotated Factor Matrix[a]

Var.	CSU	CSC	CSR	CSS	CSI	CMU	MSI	DSC	NSS	NST	NSI	EFU	ESU	ESC	ESR	ESS	EST	ESI	RAT	PAD	h^2
1.	18	14	16	18	-06	03	-04	08	11	13	-09	15	11	16	03	17	07	45	02	26	53
2.	04	17	02	-02	10	15	05	02	11	25	34	10	-10	16	20	09	12	47	12	08	61
3.	-08	27	04	13	21	08	28	-22	17	28	07	12	08	50	24	30	07	26	-03	04	88
4.	-10	48	24	23	10	18	03	01	09	20	14	05	03	38	06	18	25	16	03	17	76
5.	07	01	10	00	02	10	-07	04	12	64	03	03	13	00	00	04	-01	-08	05	-08	49
6.	21	00	23	26	19	-08	-15	-01	24	03	-09	07	02	16	07	12	16	25	-02	02	44
7.	-03	22	29	03	13	37	14	06	05	20	00	-01	03	14	10	35	24	28	-01	06	65
8.	-03	21	21	19	23	24	18	04	16	22	11	14	11	29	08	32	22	23	-08	03	69
9.	42	27	17	-02	01	25	07	-14	11	13	06	01	29	33	05	10	18	04	05	03	64
10.	-07	13	08	00	05	07	-03	03	18	28	03	03	05	39	05	22	40	25	03	10	59
11.	09	-04	23	-08	02	-11	23	13	11	03	07	28	34	12	31	04	22	21	08	-07	58
12.	53	06	-02	04	09	11	02	16	31	22	10	02	20	10	07	07	02	-06	12	-06	57
13.	20	04	24	15	-07	-12	00	-02	-01	12	19	14	01	24	01	01	-02	18	-14	-04	32
14.	13	08	31	-02	06	11	09	-03	09	-25	49	20	10	16	20	03	06	10	-02	10	60
15.	-04	13	08	04	10	06	01	09	18	15	15	62	26	04	-10	10	05	14	02	-04	62
16.	18	-09	23	25	08	15	23	18	16	13	09	11	26	17	21	27	45	08	-11	10	78
17.	-03	05	17	23	17	22	-08	19	-12	08	21	00	23	21	05	14	18	35	01	06	55
18.	-07	03	19	41	10	16	18	24	24	11	12	06	-03	16	20	09	01	21	13	05	56
19.	-04	-01	-13	11	17	-06	12	09	12	14	31	08	55	05	03	12	-02	16	05	17	60
20.	-02	-15	02	-10	05	-25	12	-11	04	15	31	06	36	-02	-14	15	05	00	-07	-02	43
21.	00	43	12	08	17	01	20	-04	13	20	06	12	06	30	22	08	18	22	09	12	60
22.	-08	26	27	04	27	07	-07	-07	18	24	10	09	15	26	16	32	28	19	06	16	71
23.	-08	41	22	09	17	07	-04	17	35	24	14	09	00	41	04	25	03	19	-08	12	79
24.	02	00	00	-15	-01	-06	60	07	09	12	33	00	21	14	02	26	09	04	02	08	67
25.	-05	21	08	09	28	12	12	05	41	15	29	00	13	25	19	10	18	27	14	-06	70
26.	-03	00	14	01	15	07	09	02	03	06	18	70	06	11	10	-01	16	12	-03	04	65
27.	00	20	14	12	-09	13	01	12	03	03	15	16	08	28	45	19	25	21	15	13	63
28.	-01	20	-05	-07	39	04	06	27	-08	07	00	10	04	-05	-09	01	17	-02	03	-09	35
29.	-02	13	46	14	04	38	01	13	01	26	09	06	-03	08	07	24	13	14	00	08	60
30.	02	13	11	14	19	18	08	-04	33	12	16	-05	-02	11	15	49	21	12	14	01	63
31.	-13	17	03	-02	03	-15	33	03	24	18	35	13	32	09	02	27	11	18	03	05	64
32.	-13	21	11	08	13	11	14	25	26	10	06	04	-01	54	-02	15	08	13	02	-06	62
33.	10	16	20	37	04	27	02	12	03	01	15	10	06	16	37	18	17	22	05	04	62
34.	28	21	07	02	05	39	-14	03	19	27	11	04	10	37	08	15	10	19	00	-18	68
35.	-03	13	02	24	30	04	03	08	-02	06	14	21	01	29	22	37	16	23	03	-06	59
36.	-02	21	07	-03	03	13	17	13	08	-03	22	24	62	09	01	16	19	11	02	11	71
37.	-12	13	26	19	22	20	-05	-05	26	19	01	04	-02	33	58	13	-05	13	-05	01	82
38.	-16	17	25	16	21	31	13	-10	25	08	-05	09	09	19	12	17	19	36	-03	-22	69
39.	-05	18	12	18	10	11	01	-06	18	10	29	-02	03	08	15	-07	31	04	10	-03	38
40.	08	03	06	-12	08	-01	06	41	20	09	-01	06	07	-05	-01	-03	24	01	07	-20	37
41.	-17	23	18	13	04	16	34	-02	18	02	12	14	06	16	03	48	-01	06	-05	-05	61
42.	06	10	18	28	-06	13	03	30	40	20	10	13	09	19	16	46	05	14	-01	17	80
43.	00	03	29	-03	24	05	-07	19	07	-03	06	03	04	33	27	19	21	23	-04	06	52
44.	34	02	26	13	13	16	01	08	11	22	16	05	-01	25	-03	20	18	12	07	04	49
45.	12	06	32	05	36	-03	-05	18	16	24	01	07	21	16	09	06	00	15	00	07	49
46.	05	20	35	21	-03	11	05	22	13	16	15	19	06	24	19	27	26	14	08	-17	67
47.	37	09	16	04	14	14	15	03	24	43	-02	06	07	26	10	14	13	08	07	08	62
48.	08	05	22	09	17	01	-04	04	02	21	-02	-05	12	12	04	15	-10	07	50	-06	46
49.	03	-15	22	07	-09	-02	03	-07	-05	-04	-13	01	09	08	01	-01	-13	-09	55	07	46
51.	-09	04	-03	01	-04	14	05	-15	04	-13	07	00	01	-02	-01	00	18	05	-03	56	43
52.	08	24	17	11	13	61	-08	-02	-01	14	-06	07	11	30	02	26	17	28	-01	02	81
53.	05	-11	26	15	-06	63	-09	-20	19	38	-09	14	18	20	06	19	23	03	-13	18	98
54.	20	15	15	-12	05	70	03	-08	25	10	06	07	-16	19	24	13	10	07	-12	11	86

[a]Decimal points omitted.

RESULTS OF THE FACTOR ANALYSIS

be avoided. The result of the fifth target-oriented rotation upon the principal-axes matrix was accepted as the final solution and is presented in Table 5 as the rotated factor matrix.

After the final rotation, differences between the unrotated and rotated communalities were no larger than .02; most of the differences were .00. Since further extensive checking seemed unnecessary, the reduced correlation matrix was not produced by post-multiplying the rotated factor matrix by its transpose. A check involving five randomly selected correlation coefficients showed that the reduced coefficients are within .02 of their respective empirical coefficients.

The interpretation of each of the 20 rotated factors is based upon the hypothesized factor content of the tests loading .30 or more upon the factor. In some interpretations to follow, the factors are defined by tests not hypothesized to cohere. In these cases, the factor interpretations are dependent not only upon the hypothesized factor contents of the tests, but also upon their newly revealed common properties.

Each factor in Table 5 is discussed in the order in which it was discussed in the section on the hypothesis. The test loadings for each successive factor will be listed along with all additional loadings (.30 or higher) of the tests, if they proved to be complex.

Interpretation of the Reference Factors

CSU Cognition of symbolic units

12. Disemvowelled Words (CSU)	.53	(.31 NSS)
9. Correct Spelling (ESU)	.42	(.33 ESC)
47. Word Transformation (NST)	.37	(.43 NST)
44. Word Combinations (CSU)	.34	

Disemvowelled Words, although shortened, once again leads the tests that are loaded on the CSU factor. It would appear from the test's history that the factor is concerned with the recognition of complete and correct words; a factor that might be called "word closure" (Pemberton, 1953).

Such a conclusion is strengthened by the fact that Correct Spelling, designed as a measure of ESU, is also a measure of the recognition (perhaps not in a critical light) of complete and correctly spelled words. Correct Spelling was hypothesized to be a test of ESU because sensitivity to the correctness or incorrectness of commonly misspelled words was thought to be evaluative. Possible reasons for its being a CSU test instead will be offered in the discussion of factor ESU.

Word Transformation's loading on the CSU factor implies that the careful analysis of previous test data did not accomplish all that it intended. Tests for both CSU and NST were correlated in the previous analysis (Guilford, et al., 1961) in which those factors were discovered. From Word Transformation's correlation with the CSU factor, although secondary, it appears that the symbolic redefinition task is dependent upon recognition of the symbolic units needed in effecting the transformation. Word Combinations once again has its primary loading on CSU with no significant loadings on the NST factor for which it had been originally designed.

CSC Cognition of symbolic classes

4. Best Number Pairs (ESC)	.48	(.38 ESC)
21. Number Classification (CSC)	.43	(.30 ESC)
23. Number-Group Naming (CSC)	.41	(.41 ESC; .35 NSS)

The two shortened forms of the tests selected to measure the CSC factor function as anticipated, except that the tests for CSC are led by a test designed for ESC. Like its analogous semantic test, Best Word Pairs (Nihira, et al., 1964), Best Number Pairs contributes more to the cognition-of-classes factor than to the evaluation factor. Like Best Word Pairs, Best Number Pairs asks E to evaluate which pair of stimuli makes the best class, but the specific properties of the best class, in other words, the specific criteria, are not defined in each item. Since the specific aspects of the best class need to be discovered for each item, cognition abilities should be expected to determine much of the test's variance. The CSC factor defined in this analysis could be confined to the ability to recognize common properties of numbers, but a letter test has previously been loaded on it (Guilford, et al., 1961).

The significant side loadings of the three CSC tests on the ESC factor impressively demonstrate that tests of the two factors are highly related. Although it could be contended that no separation between CSC and ESC tests should have been attempted in rotating axes, it was felt that an unclear separation would be heuristically more valuable than the report of a less clearly defined, composite, classes factor.

CSR Cognition of symbolic relations

29. Seeing Trends II (CSR)	.46	(.38 CMU)
46. Word Relations (CSR)	.35	
45. Word Patterns (CSI)	.32	(.36 CSI)
14. Form Reasoning (NSI)	.31	(.49 NSI)

Both CSR tests function as anticipated in this analysis. Although they had undergone considerable shortening, it appears that their variances remain dominated by CSR. The significant CMU side loading for Seeing Trends II is not reasonable as the trends are based solely on the letter content of the words, and not on their meanings. The result is consistent with a similar CMU side loading found for Seeing Trends II by Guilford, et al. (1961). Some semantic recognition must somehow be involved in the test. Although recognition of the symbolic trends is indicated by verbal statement of E, the scoring procedure did not weight, but in a minimum manner, E's verbal abilities at expressing the symbolic trend.

Word Patterns and Form Reasoning, measures of CSI and NSI respectively, have small but significant side loadings on CSR. This is not the first time that a separation between relations tests and implications tests has been difficult to achieve (see Gershon, et al., 1963). Each of these tests did have its primary loading on the factor for which it was hypothesized.

CSS Cognition of symbolic systems

18. Letter Triangle (CSS)	.41	
33. Similar Pairs (ESR)	.37	(.37 ESR)
6. Circle Reasoning (CSS)	.26	

Letter Triangle leads the factor called CSS in this analysis. Although both Letter Triangle and Circle Reasoning defined CSS factors with moderate loadings in several previous investigations (Green, et al., 1953; Guilford, et al., 1954; Kettner, et al., 1959; Guilford, et al., 1961), they did not emerge together in a more recent investigation by Petersen, et al. (1963). The implication derived from the unsuccessful attempts to isolate a clear CSS doublet is that additional hypothesized measures of CSS should be developed or refined and analyzed in the future so that several measures would be available that consistently cohere.

Similar Pairs' loading on the CSS factor should be accounted for by cognition that might be necessary to recognize the relations within word pairs. The determination of these relations is dependent upon alphabetical order, which also determines the systems in Letter Triangle, and upon need for letter rearrangements in some items.

CSI Cognition of symbolic implications

28. S Test (ESI)	.39	
45. Word Patterns (CSI)	.36	(.32 CSR)
35. Symbol Grouping (CSI)	.30	(.37 ESS)

The ability to foresee symbolic implications, CSI, is defined in this analysis by the two tests designed to measure the factor, but is measured best by the S Test, a test originally designed to measure sensitivity to problems. The original F Test, from which the S Test was

derived, probably did not indicate the "sensitivity-to-problems" factor because the factor was semantic. The semantic tests that did bring out the sensitivity-to-problems factor, were later shown to belong largely on factor CMI, the cognition ability parallel to CSI, on which we now see the S Test loaded. It appears that there is no difference between the ability to see symbolic implications and being sensitive to them.

The two CSI tests, Word Patterns and Symbol Grouping, although shortened to make them more univocal, still appear to be complex. Symbol Grouping's primary loading on ESS appears to reflect the systems (ordering) aspect of the items; some items require relatively little foresight and more systematic, almost rote, grouping of symbols. Evaluation probably enters because, in attempting to achieve the most efficient solution, E needs to compare one solution with another and decide which is more efficient.

CMU Cognition of semantic units

54. SCAT Verbal (CMU)	.70	
53. PSAT Verbal (CMU)	.63	(.38 NST)
52. ITED General Vocabulary (CMU)	.61	(.30 ESC)
34. Sound Grouping (ESC)	.39	(.37 ESC)
29. Seeing Trends II (CSR)	.38	(.46 CSR)
7. Correct Letter Orders (ESS)	.37	(.35 ESS)
38. Symbol Reasoning (ESI)	.31	(.36 ESI)

The verbal-comprehension factor is clearly defined by the three tests selected to measure CMU. The tests are relatively univocal, having very high loadings on CMU and small or no loadings on symbolic factors. The tendency of the CMU tests to be loaded about .20 on many of the remaining factors in the analysis probably is accounted for by variance introduced in the understanding of instructions or in verbalizing aspects of solving symbolic problems.

Once again, Sound Grouping demonstrates its factorial complexity by splitting its variance between ESC and CMU. Familiarity with the words used as stimuli for this test appears to facilitate either the pronunciation or the classification based upon the pronunciations. One might expect an analogous phenomenon underlying the CMU loading of Seeing Trends II. But knowledge of meanings of the words in Seeing Trends II in no way aids in the discovery of the symbolic trend. The only common element in the CMU tests and Seeing Trends II is that words are employed.

Correct Letter Order's CMU loading is rationally tenuous, the only CMU variance possibly being contributed by the verbal descriptions of the letter series. Symbol Reasoning's almost negligible CMU loading appears to be fortuitous, as no items employ words or verbally complex criteria. It seems possible, however, that Es may translate the symbolic syllogisms into some verbal form and then perform the required task.

MSI Memory for symbolic implications

24. Numerical Operations (MSI)	.60	(.33 NSI)
41. Way-Out Numbers (ESS)	.34	(.48 ESS)
31. Sign Changes (NSI)	.33	(.35 NSI; .32 ESU)

Numerical Operations emerges clearly as the test defining the MSI factor. MSI receives further support from Way-Out Numbers due to the errors in item construction mentioned earlier, and from Sign Changes, a test designed to measure NSI, but which has consistently shared MSI variance (Petersen, et al., 1963). It appears that the MSI-NSI sharing is reciprocal; while Sign Changes is loaded slightly on MSI, Numerical Operations is loaded slightly on NSI. The shared variance on these two factors is not great enough, however, to cause serious concern regarding their distinctness.

DSC Divergent production of symbolic classes

40. Varied Symbols (DSC)	.41	
42. Word Changes (NSS)	.30	(.40 NSS; .46 ESS)

Varied Symbols once again serves as the principal measure of the DSC factor. Number Grouping, which was not loaded on DSC in a previous investigation (Gershon, et al., 1963), failed to be again. The strong face validity of Number Grouping as a measure of DSC apparently is misleading as the test loaded on the ESS factor in this analysis. No explanation could be found for the small tertiary DSC loading of Word Changes other than that Es may employ a discursive trial-and-error ordering procedure which would mean divergent-thinking processes.

NSS Convergent production of symbolic systems

25. Operations Sequence (NSS)	.41	
42. Word Changes (NSS)	.40	(.46 ESS; .30 DSC)
23. Number-Group Naming (CSC)	.35	(.41 CSC; .41 ESC)
30. Series Relations (ESS)	.33	(.49 ESS)
12. Disemvowelled Words (CSU)	.31	(.53 CSU)

The two tests selected to measure NSS perform again in a reliable manner. It appears that Operations Sequence is one of the strongest and is the most univocal measure of this factor. The large side loading of Word Changes on ESS might be due to a strategy that employs alternate orderings that are quickly evaluated according to the limitations imposed by criteria given in the test instructions. A similar rationale could explain the minor secondary NSS loading of Series Relations. E tries out each given operation in turn, producing a fully determined series.

The small but significant loadings of Number-Group Naming and Disemvowelled Words on NSS cannot be explained, as neither test appears to involve system production.

NST Convergent production of symbolic transformations

5. Camouflaged Words (NST)	.64	
47. Word Transformation (NST)	.43	(.37 CSU)
53. PSAT Verbal (CMU)	.38	(.63 CMU)

The NST factor emerged in this analysis with both tests selected to measure it loading primarily upon it. The CSU side loading of Word Transformations was discussed in connection with the CSU factor. From the PSAT Verbal test's loading on NST, it appears that NST may play a role in responding to verbal tests — the emergence of words or word roots may aid in the comprehension of word meanings and complex ideas.

NSI Convergent production of symbolic implications

14. Form Reasoning (NSI)	**.49**	**(.31 CSR)**
31. Sign Changes (NSI)	**.35**	**(.33 MSI; .32 ESU)**
2. Best Letter Set (ESS)	**.34**	**(.47 ESI)**
24. Numerical Operations (MSI)	**.33**	**(.60 MSI)**
20. Marking Speed (MOT)	**.31**	**(.36 ESU)**
19. Letter "U" (ESU)	**.31**	**(.55 ESU)**

The two tests selected to measure NSI perform in this analysis as expected, except that both have some complexity. In addition to the NSI tests' being complex, the NSI factor found in this analysis shows minor loadings in a few heterogeneous types of tests. The factor is defined in part by an MSI test that has previously shared NSI variance (Petersen, et al., 1963), and had been actually thought to be an NSI measure for a long time. Shifting it primarily to the MSI factor resulted from finding that it is strongly related to the Wechsler Digit Symbol test, which is more obviously a memory test (Davis, 1956; deMille, 1962). In addition, Marking Speed and Letter "U", both of which are speeded tests, share variance with NSI, apparently due to their common speeded natures. Had the rotational solution insisted upon a separate speed factor, with Marking Speed the leading variable on it, this test would probably not have appeared in this list, or in the list for factor ESU to be presented a bit later. Best Letter Set's loading on NSI cannot be logically explained by common aspects of the tests; there appear to be none.

To date, this solution represents the clearest separation between the MSI and NSI factors, although there still seems to be some common aspect in the tests used here to measure them and more univocal tests are apparently needed for both.

EFU Evaluation of figural units

26. Perceptual Speed (EFU)	**.70**
15. Identical Forms (EFU)	**.62**

The two EFU tests perform exactly as hypothesized, being univocal and highly saturated with common-factor variance. The EFU factor is the clearest interpretable factor to emerge in this analysis, probably due to its relative dissimilarity of content from the symbolic and semantic factors.

Interpretation of the Experimental Factors

ESU Evaluation of symbolic units

36. Symbol Identities (ESU)	**.62**	
19. Letter "U" (ESU)	**.55**	**(.31 NSI)**
20. Marking Speed (MOT)	**.36**	**(.31 NSI)**
11. Derivations (ESU)	**.34**	**(.31 ESR)**
31. Sign Changes (NSI)	**.32**	**(.35 NSI; .33 MSI)**

The ESU factor emerged with remarkable clarity. The two speeded sensitivity tests that were previously suggested as measures of ESU (Guilford and Hoepfner, 1963) lead the factor with high loadings and little or no complexity. It is interesting to note that with strong tests of EFU and a sufficient number of tests for ESU, including Symbol Identities and Letter "U", in the analysis, the two factors separate very clearly. This decisive result clears up earlier uncertainties as to whether "perceptual-speed" tests composed of literal material should go with Thurstone's original perceptual factor or should represent a separate factor (Thurstone, 1938b; Coombs, 1941; Bechtoldt, 1947).

The third ESU test loaded on the ESU factor is Derivations, also a sensitivity test, involving words. Based upon the three ESU tests loading on this factor, it appears that ESU is the ability to make rapid decisions regarding the symbolic identity or accuracy of words, letter sets, and number sets. In Symbol Identities, there is a comparison of two given symbolic units to determine whether or not they are identical. In Letter "U", a word class is specified (words containing the letter "U"), with E to say whether or not each word satisfies the specification. Symbol Identities is a direct parallel to figural tests of EFU, in which figures are to be compared, with E to decide whether or not they contain exactly the same elements. Letter "U" is parallel in some important respects with Double Descriptions (Nihira, et al., 1964). In that test, E is given a class specification for a verbalized concept, e.g., things both hard and round, and he must say which of four offered semantic units best fulfills the specifications. Besides the difference in kind of information, which determines two distinct abilities, there is another difference, which probably is not very important, and that is the yes-no type of decision (sensitivity) in Letter "U" and the relative decision (estimation) in Double Description.

Derivations does not fit either of the two item models just described for Symbol Identities and Letter "U". The things being compared are not exactly the same except for one element, as in the former, nor are class specifications given, as in the latter. The letters of the short word said to be extracted from the long word must coincide with a completely identical set of letters in the long word, except that the order is probably different. There is no clear model presented for comparison. This may be a reason for the lower ESU loading for Derivations than for Symbol Identities.

The two ESU "misses," those tests hypothesized for ESU but not loaded significantly on it (Correct Spelling and Familiar Letter Combinations), also aid in interpreting the ESU factor by indicating what ESU is not. One characteristic the two "misses" have in common but do not share with the other three ESU tests, is that the things with which comparisons must be made are not given on the printed page. They can only be compared with something in memory storage or something retrieved from memory storage, perhaps in the form of an image. In Correct Spelling, the needed model is the remembered correct spelling of each word. Some of these models E would have and some he would not. The task boils down to the question of how many of the 120 words in the test does E know, spelled correctly. This statement of the task makes it appear like a measure of cognition, as it turned out by analysis to be.

In Familiar Letter Combinations, the criterion for judgment is familiarity of the syllables or their observed probability of occurrence in E's experience. In this test, there is no clear model for E to use, and what he has to use is also something in or from memory storage. Although it appears that the memory feature applies especially to the two ESU tests that missed, the question arises as to how general the implied evaluation principle is. If it is quite general, the principles

that comparisons must be between perceived information would place an important restriction upon the definition of evaluation abilities.

The presence of Marking Speed, along with Sign Changes, here suggests some confounding of a finger-speed factor with both NSI and ESU. As was suggested in the discussion of NSI, the rotation to an additional finger-speed factor might have cleared up the picture for both NSI and ESU. Implied is a general principle, namely, that the appearance of obliqueness among factors may be due to lack of a sufficient number of dimensions being included in orthogonal rotations.

ESC Evaluation of symbolic classes

32. Sign Changes II (ESR)	.54	
3. Best Number Class (ESC)	.50	(.30 ESS)
23. Number-Group Naming (CSC)	.41	(.41 CSC; .35 NSS)
10. Decoding (EST)	.39	(.40 EST)
4. Best Number Pairs (ESC)	.38	(.48 CSC)
34. Sound Grouping (ESC)	.37	(.39 CMU)
43. Word Choice (ESC)	.33	
37. Symbol Manipulation (ESR)	.33	(.58 ESR)
9. Correct Spelling (ESU)	.33	(.42 CSU)
21. Number Classification (CSC)	.30	(.43 CSC)
52. ITED General Vocabulary (CMU)	.30	(.61 CMU)

Although all four tests designed to be measures of the ESC factor are loaded on the factor called ESC, the factor is the least clear of all the new experimental factors found. Three of the ESC tests have complexities of two, and the factor is led by a test designed for ESR. Usually, inspection of the common characteristics of the tests leading a factor reveals the essential nature of the common aspect that the factor represents. In the case of the ESC factor, the common element of the leading tests is rather obscure.

Sign Changes II requires the examinee to change signs in a numerical expression so that the expression becomes an equation. Introspectively, it seems that a successful attack on such problems would include the tactic of becoming aware of what both sides of the expression have potentially in common, and from this awareness, to make the appropriate sign changes to bring about that common numerical value and thus change the expression into an equation. To clarify this attack with an example, consider the sample item: $3 + 1 = 6 \times 2$. The first step in effectively changing this expression into an equation is not to substitute signs, but to be sensitive to what the pair of numerical value 3 and 1, and the pair 6 and 2 potentially have in common. Their common, or class property is either the numerical value of 4 ($3 + 1$ and $6 - 2$), or 3 (3×1 and $6 \div 2$). Since only one of the necessary sign changes is given as an alternative answer, the only acceptable solution is the first one, and the common element in the expression is the value of 4, the value for which the signs must be changed. But this line of thinking suggests cognition rather than evaluation, factor CSC rather than ESC, and the loading for this test on CSC is only .17. Another hypothesis is that E somehow takes the offered solutions as classes of operation changes and considers them for adequacy.

The task involved in Best Number Class is more definitely to realize the numerical classifications of given numbers and then to select the one classification that is most valuable, value of number classes being defined by the test as the criterion.

The test, Decoding, designed as a measure of EST, also is loaded equally on the ESC factor. The reason for its loading is probably due to the class-properties variance that is introduced in E's analysis of the words into classes according to the rules for the codes, each rule representing a different class property. It is necessary to recognize the parts of the uncoded words as belonging to certain classes of codability and then to make decisions as to the values of the recognized classes.

The three ESC tests with the smaller significant loadings on this factor were constructed as estimation tests of ESC. These tests, except for Best Number Pairs, involve making a choice among given class properties on the basis of criteria supplied in the test. Although it would seem reasonable to have rotated the ESC axis to maximize its correlations with these more simple ESC tests, their tendency toward factor complexity and the resulting loss of simple structure weighed against such a move.

The loadings of Symbol Manipulation, Correct Spelling, and ITED General Vocabulary on the ESC factor are small and not important to the interpretation of this factor when the distinctly larger loadings on their respective hypothesized factors are considered.

ESR Evaluation of symbolic relations

37. Symbol Manipulation (ESR)	.58	(.33 ESC)
27. Related Words I (ESR)	.45	
33. Similar Pairs (ESR)	.37	(.37 CSS)
11. Derivations (ESU)	.31	(.34 ESU)

Three of the four tests designed to measure ESR are loaded on the ESR factor, clearly defining it as representing the ability to make choices among symbolic relationships on the basis of identity and consistency. The significant side loadings of the ESR tests were mentioned before in connection with the respective factors upon which the complex ESR tests are loaded—factors ESC and CSS.

Derivations' small loading on ESR could be due to the whole-part relations between the units in the test. Short words were to be judged as being parts of longer words or not being parts. When small or simple words are to be judged as derived or not derived from the given word, a choice based upon such a part-whole relationship is probably more efficient than a comparison of the elements in the words for unit identities.

ESS Evaluation of symbolic systems

30. Series Relations (ESS)	.49	(.33 NSS)
41. Way-Out Numbers (ESS)	.48	(.34 MSI)
42. Word Changes (NSS)	.46	(.40 NSS; .30 DSC)
35. Symbol Grouping (CSI)	.37	(.30 CSI)
7. Correct Letter Orders (ESS)	.35	(.37 CMU)
8. Correct Number Series (ESS)	.32	
22. Number Grouping (DSC)	.32	

Four of the five tests designed for ESS came out significantly loaded on ESS. The two leading tests, Series Relations and Way-Out Numbers, are in the estimation category and are composed of numbers. The two with distinctly smaller loadings, Correct Letter Orders and Correct Number Series, are in the sensitivity category, one being a number test and the other a letter test. It may be of some interest that while

sensitivity tests proved to be better for ESU, estimation tests proved better for ESS. The trend must be better supported, however, before a principle can be stated. It is probably significant, however, that a system can deviate from a standard of comparison much more readily by degrees than can a unit.

Series Relations and Way-Out Numbers differ somewhat in the operations that they require. In the former, E probably makes new systems (series), following rules given in the alternative responses. He then compares each new system with the model that is given, deciding which new one comes nearest. The criterion is the degree of closeness of one set of numbers to another set. In the latter test, he is virtually to compare the distances of the first and last numbers in an irregular series to decide which one is farther from the other numbers. It is as if he were treating the same series first as one system and then as another, or from one point of view then from another. The criterion is numerical distance. The difference between these two relatively strong ESS tests contributes some breadth of the nature of the factor.

The two weaker ESS tests are close to being alternate forms of the same test, and both differ from the two stronger tests, as indicated before. They present letter series in the one case and number series in the other, with a verbal description of the principle that should be satisfied in a series. Sometimes the series follows the principle exactly, sometimes not. The sample for a letter series given in the Appendix describes the series thus: "Alternate letters in the alphabet (skipping one)." The sample for a number series is: "Alternately add 1, multiply by 3." It is probably significant that Correct Letter Orders has a significant loading on CMU whereas Correct Number Series does not. The illustrated rules are somewhat, although not universally, typical. But it should matter more whether verbal terms are correctly understood in the first example than in the second. In the latter the rule can be more simply and precisely stated.

Word Changes' primary loading on ESS was explained earlier as possibly resulting from a strategy of rapidly ordering the given words and then evaluating the orders to determine whether they are consistent with the rules of the test. In this analysis, the evaluation variance proved to be fully as strong as the convergent-production variance. In a similar manner, Symbol Grouping's ESS loading was explained as possibly resulting from evaluating trial orderings of the symbols to determine whether they are consistent with the rules and are most efficient for achieving the final, accepted ordering. This kind of result suggests the generalization that a number of cognitive and convergent-production tests may actually involve some evaluation variance which has heretofore been missed because they have not been analyzed along with definitive evaluation tests.

Number Grouping's only significant loading is on ESS. No compelling reason is apparent for this loading, unless divergent-production tests, too, may become involved with evaluation variance. But there is generally much doubt about Number Grouping, for it was not loaded on DSC in its previous factor-analytic investigation and appears not to be clearly a measure of any factor found in the analyses in which it has been employed.

EST Evaluation of symbolic transformations

16. Jumbled Words (EST)	**.45**	
10. Decoding (EST)	**.40**	**(.39 ESC)**
39. Typing Errors (EST)	**.31**	

The three tests designed to measure EST emerged on the EST factor with unexpected clarity. The construction of the EST tests had proved to be most difficult and therefore it was expected that EST might not be found in this analysis. All three tests designed for EST, and only those three, came out significantly on EST. All three tests involve the use of words. The strongest for EST was considered to be a sensitivity test, the other two were considered estimation tests, but one of them is not far behind the leading test, so there is no conclusion except that it probably does not matter which type is used for EST measurement.

The three tests differ somewhat in terms of operations that E probably performs as he takes the tests. In Jumbled Words, he decides whether each response word could have come from the given word merely by rearrangement of the letters it contains. The criterion is in terms of a certain invariance or of element identity under the transformation of changed order.

In Decoding, E is expected to apply certain rules, of which five are given, in coding the letters of each of two words into a sequence of digits. The transformation in each case is in the form of substitution of elements according to rules. E then is expected to decide which coding (transformation) result could be most easily and correctly decoded. The difference in ease and correctness of decoding depends upon the approach of the substitutions for a word to univocality, under the rules. That is, some rules pertain to a smaller class of letters. For example, take the two rules: For all double letters (oo, gg, etc.), substitute the digit 1; for each single consonant, substitute the digit 5. There being more alternatives among single consonants than among occurrences of double letters, a 1 is more indicative in a coded word than is a 5, for which uncertainty in decoding would be greater. Assuming that E follows this principle, the criterion is the preponderance of high-probability digits in each word. E has a third category of response that he uses if he thinks the two coded words are equally near univocality.

In Typing Errors, E is presented with what he is told is a misspelled word, e.g., "WORM." His alternative answers are: "WARM," "WORN," and "WARS." From a knowledge of the arrangement of letters on the keyboard of a typewriter, which transformation in spelling (substitution of one letter element) has most probably occurred? The decision boils down to which letter change involves two letters closest together on the keyboard. Incidentally, the keyboard is printed on each page of the test booklet. A correlation of scores on this test with amount of typing experience was very low, indicating that typing experience contributed little or nothing to variance in the scores. A judgment of nearness of two letters on the keyboard may involve some other ability, perhaps distance estimation, which may account for the minimally significant loading on EST for this test. Whatever EST variance there is, may come from the comparison of alternative transformations in the form of letter substitutions.

To summarize these comparisons, the transformation was in the form of reordering of letters in one test and in the form of letter substitutions and letter-digit substitutions in the other two tests. The criteria appear to be identity of elements in spite of transformation in one test, univocality of reference in the coding test, and nearness of position in a symbolic system (keyboard) in the third. These differences suggest that there is some scope in kinds of transformations and in criteria for decision involved in connection with factor EST.

ESI Evaluation of symbolic implications

2. Best Letter Set (ESS)	**.47**	**(.34 NSI)**
1. Abbreviations (ESI)	**.45**	
38. Symbol Reasoning (ESI)	**.36**	**(.31 CMU)**
17. Letter Problems (ESI)	**.35**	

The ESI factor is defined by three tests designed as measures of ESI, but is led by a test designed for ESS that is not loaded on that factor. The most obvious common characteristic of the two leading tests is that the given stimuli are sets of letters, but this is also true of many other tests in the battery. But in addition, the item formats of the two tests are similar (see Appendix). In Abbreviations, a sequence of three or four letters is presented as the potential abbreviation of three alternative, familiar words, with E to say for which word the abbreviation best stands or that it best implies. In Best Letter Set, a sequence of three or four letters is given and three alternative letter sets of the same length, E to select the one that is most similar to the given set by virtue of common properties. The test was expected to be a measure of ESS because it was thought that the nearness of one set to another in terms of their constitutions would be a matter of comparing systems for approach to identity in terms of systemic properties. The given set might be W V U, which is easily recognized as three consecutive letters in reverse alphabetical order. The three alternative answers are: D C A, T P O, and Z P E. None of these is exactly in reverse consecutive order. The first two are in reverse order, but the very first one comes nearer to consecutive order.

It is not very clear how Best Letter Set becomes an implications test rather than a systems test. It is little more than stating the obvious to say that the given set apparently implies the best alternative, similarly to the way in which a letter set in Abbreviations implies a longer letter set in the form of a real word. A revised format of Best Letter Set, giving alternatives that either do or do not fit the principle of the given set exactly, might have been a better ESS test. But there would still be much unanswered about underlying reasons for the difference in factor content of the two test formats.

Symbol Reasoning would seem to be an ideal type of test for ESI, for it requires E to decide whether conclusions, expressed in symbolic form, can or cannot be justifiably drawn from other symbolic statements in the form of equations or inequalities. In essence, this test would seem parallel to the verbal-syllogistic test Logical Reasoning, which has more recently been found to measure factor EMI (Nihira, et al., 1964). Just as Logical Reasoning proved to be somewhat complex factorially, it is probable that Symbol Reasoning is complex, as suggested by its low loading of .36 and its secondary loading of .31 on CMU. The latter component may mean that there is some verbalizing of the problems or there is some difficulty in understanding the instructions, which are long and somewhat involved. A much simpler task is needed in order to achieve a more univocal test. One would have to take care, however, to steer away from factor ESR, for which a similar test, Symbol Manipulation, is the leading test. The parallel test Logical Reasoning shares some variance with factor EMR. It may be that a univocal test for ESI lies in a little different direction than a test like Symbol Reasoning. Of the tests in the list above, Abbreviations, of course, appears to hold the most promise.

Letter Problems holds some promise of univocality, so far as this analysis goes, but with a low loading of .35 and a reliability of .88, there is considerable room for additional common-factor content. Letter Problems is interesting for the fact that it was developed by analogy to Form Reasoning (see the Appendix), which is a marker test for factor NSI. A major difference, which should not be factorially significant, is that Letter Problems uses letters as symbolic elements whereas Form Reasoning uses geometric forms as symbolic elements. The significant difference is that Form Reasoning calls for inferences or conclusions to equations of a certain type whereas Letter Problems asks for decisions as to whether a problem equation is solvable (has a determined conclusion), is not solvable, or is solvable with a minor change in the problem. It might be said that the criterion for evaluation is solvability or the possibility of valid inferences or implications. This is not strictly a matter of judging the value or identity of an implication, as such, or its similarity to another implication, which are common kinds of criteria in tests of other evaluation factors. But if a new kind of criterion is involved and is crucial to the loading on ESI, we have a little extension of connotation of evaluation abilities and the evaluative process as envisaged from the psychometric approach.

Interpretation of the Experimental Motivation Factors

The last two factors to be reported are each confined to measures of a particular kind and represent a failure to obtain a somewhat general motivation factor thought to underlie all the measures. A lack of correlation among some of the four hypothesized motivation measures made it impossible to force the measures onto one factor.

RAT Ratings

49. Rating - Booklet Effort (MOT)	**.55**
48. Rating - Test Liking (MOT)	**.50**

The tendency for self-ratings to reflect bias, and thus intercorrelate due to personal consistency in bias, probably accounts for most of the specific nature of this factor. Similar formats of rating and similar tasks also contribute to the specific nature. Whatever this factor represents, it is clear that it is not related to the aptitude measures in the battery.

PAD Prediction-achievement discrepancy

51. Prediction-Achievement Discrepancy (MOT)	**.56**

This singlet is the second example of failure to obtain the hypothesized motivation factor. The least-squares discrepancy from the regression of academic achievement on DAT subscores appears to be not so much due to common-factor motivational aspects of the examinees as it is due to error in the predictors, criterion, or other traits that contribute to grade-point averages but are not measured by the nine DAT subtests. If the discrepancy were due to motivational factors, the scores would presumably correlate more highly with the other measures that presumably contain some motivational variance.

The fact that certain aptitude scores were subtracted to obtain the PAD score is the reason why variable 51 correlates low or negatively with all the aptitude measures, a number of which are related to the DAT scores that were subtracted. Having no large correlations with any variate, but small negative ones with most of them, the PAD score formed a singlet factor of its own, which is partially or wholly spurious.

DISCUSSION

In this section, the experimental findings of this study are considered in terms of how they bear upon the hypothesis with which the study was initiated, both in terms of how they indicate the continued fruitfulness of the structure-of-intellect theory as a source of testable hypotheses and more specifically how well types of abilities in the model were differentiated. Relations of the findings with respect to symbolic-evaluation factors to the very limited prior indications of any of those factors will be noted, as well as with respect to the findings of a recent parallel study of semantic-evaluation abilities. General light that is thrown upon the nature of evaluation as a category of abilities and as a distinct kind of operation will be considered. Ending the section will be the usual recommendations regarding tests for the factors under special investigation.

Relations of Results to the Hypothesis

Since so many factors pertaining to the structure-of-intellect model were under examination along with the symbolic-evaluation set of factors, a somewhat general consideration can be given to how well factor-analytic operations of the kinds used can differentiate factors in different categories. Abilities will be distinguished in terms of operations, in terms of contents, then in terms of products, with special emphasis upon the experimental evaluation factors.

Relation to Hypothesized Operations

All of the five kinds of operation were involved in this study, but obviously not equally so. There were five factors of symbolic cognition, the sixth one (CST) not yet having been demonstrated, and one of semantic cognition (CMU); one memory factor (MSI); one divergent-production factor (DSC); three convergent-production factors (NSS, NST, and NSI); and seven evaluation factors, one figural (EFU) and the six symbolic-evaluation factors of special interest. The factors outside the last-mentioned six were reference factors, with the usual concern lest some of the new factors not be shown to be distinct from parallel factors in other operation categories or lest some of the variances of new experimental tests not be well accounted for. The greatest concern was regarding the demonstration that the symbolic-evaluation abilities be differentiated from the parallel cognition factors. In constructing evaluation tests, it is not easy to be sure that cognitive variance has been ruled out by the experimental controls in the test conditions, or even that cognitive variance may not dominate the test. Other reference factors were brought in because new experimental tests were suspected of involving still other operation variances.

The general outcome was that six symbolic-evaluation abilities could be demonstrated as distinct from abilities of cognition, memory, and divergent and convergent production. How well this was accomplished is indicated by the fact that 22 of 25 experimental symbolic-evaluation tests cohered with their respective factors in that category.

Another, more general, indicator is the frequency with which tests designed for the various operations have significant loadings on factors in those operation categories, although this does not tell the whole story. On the factors of cognition (CSU, CSC, CSR, CSS, CSI, and CMU) there were 15 such loadings for cognition tests, two for convergent-production tests, and seven for evaluation tests. As feared, there was greatest danger that evaluation tests would be involved with cognition variances. But the picture is somewhat better when we consider that there were only 13 tests thought in advance to be cognitive (two had two loadings each on cognitive factors, making 15 loadings) and there were 27 tests slated for evaluation.

With memory abilities and divergent-production abilities represented by one factor each, it is not possible to extract anything by way of generalization about those two operation categories. As to tests loaded significantly on the three convergent-production factors (NSS, NST, and NSI), all six of the convergent-production tests had significant loadings on their respective factors. There were also three from the category of cognition, one from memory, and three from evaluation.

On the seven evaluation factors there were 28 significant loadings for evaluation tests, two for cognition tests, one for a divergent-production test, and two for a convergent-production test. It would appear to be easier to keep evaluation out of cognition tests than to keep cognition out of evaluation tests, although this depends somewhat upon where the rotations happen to go in a particular analysis. At any rate, from another point of view the six experimental evaluation factors appear to be well differentiated from other operation factors. Although one might expect evaluation to play roles in connection with convergent-production tests, in which responses must be narrowed down to a single one for each item, as far as can be seen there is little reason to doubt the distinctness of those two operation categories.

Relation to Hypothesized Contents

There being only one figural factor and one semantic factor in the analysis, it is not easy to derive

much in the way of a test of the hypothesis of three distinct kinds of abilities as to content or kind of information. But in the graphic rotations, it was very apparent that the figural and semantic factors were separate from the remaining symbolic factors. Of the 16 symbolic factors, none had on them any significant loadings from the two figural-evaluation tests and two had significant but small loadings, one each, from two semantic tests. For the one semantic factor (CMU), four of the seven significant loadings were in symbolic tests. The latter kind of result has been rationalized on the basis of CMU variance contributed from variations in verbal understanding of the instructions and/or from E's verbalizing his work on the items of the four tests.

Another point regarding kind of information should be mentioned. This pertains to what kinds of test materials can be used as stimuli for eliciting symbolic information. The results of this study demonstrate quite clearly that symbols can be in the form of numbers, letters, syllables, words, or even geometric forms serving as signs or tokens. No experimental (symbolic-evaluation) factor is defined solely by tests employing one kind of symbol to the exclusion of the others. Instead, most of the factors are defined by all three types of symbols or by tests employing combinations of the types.

Relation to Hypothesized Products

Among the six hypothesized symbolic-evaluation factors, there is very clear delineation. From inspection of the significant loadings, there appears to be very little sign of obliqueness between any two of the six factors. The factors ESU, ESS, and EST show no significant loadings from tests of other symbolic-evaluation factors. Factors ESR and ESI show one such loading each — ESR for an ESU test with a minimal loading of .31, but ESI with an ESS test taking the lead for ESI with a loading of .47. But neither result throws any doubt upon the distinctness of factors. Factor ESC shows the most involvement in tests for other ES factors, having its highest loading of .54 from an (expected) ESR test and a very small loading of .33 for another (actual) ESR test. ESR does not reciprocate by having any significant loadings for ESC tests. Other "foreign" tests loaded on ESC are an ESU test with .33 and an EST test with .39. Other sets of six tests in a column of the model have shown less clarity of separation than is indicated among these six evaluation factors.

Relations to Previous Results on Evaluation Factors

As mentioned in the discussion of the previous literature, there has been very little clear precedent for any of the six experimental factors of this study. There were tenuous evidences for only two of them — factors ESU and ESR. As reported by French (1951), Thurstone, Coombs, Bechtoldt, and others sometimes found tests calling for the identification of letters or the matching of letters, and even combinations of letters, sometimes to appear loaded on perceptual speed (now identified as EFU) and sometimes suggesting a separate factor (a possible ESU). The trouble was that a distinction between figural and symbolic information had not been recognized and not enough tests of the two kinds were always analyzed together. Letters taken singly are likely to serve as material for figural information, and cancellation tests and the Scattered X's test are more definitely more figural than symbolic. The Letter A test, which is essentially an alternate form of Letter "U" used in this study, usually had a weak EFU loading and tended to help suggest the second factor. This is the only link between a possibly earlier ESU and the factor demonstrated in this study.

The possible earlier sign of a factor ESR came out of analyses at the Aptitudes Research Project at the University of Southern California, in which Symbol Manipulation tended to lead on a factor tentatively placed in the unique spot ESR because Symbol Manipulation is a kind of symbolic-syllogism test, and a verbal-syllogism test consistently defined a factor called "logical evaluation," and in due time it was placed in the parallel cell of EMR. Symbol Manipulation also sometimes had significant loadings on the logical-evaluation factor, which helped to support its factor's claim to the cell ESR. From the liberal support given to the interpretation of the ESR factor in this study, it appears that the symbol manipulation factor was properly placed, and the test by that name still leads on the factor by a good margin.

A comparison of the results of this study with those from the recently reported analysis of semantic-evaluation factors (Nihira, et al., 1964) is of some interest. Although many of the tests are different in form, there are naturally some parallels. For example, tests of ESC were direct translations of the EMC tests into symbolic content. Best Number Pairs was an adaptation of Best Word Pairs. The two performed very similarly in that both went rather on their respective cognition parallels, factors CSC and CMC. In both cases it can be pointed out that for measurement of evaluation rather than cognition, there is a need to state explicitly the criteria upon which judgments are to be made, and possibly provide models for comparison or to describe them, as in Double Descriptions and Letter "U".

Another parallel between results in the two studies concerns the implications factors ESI and EMI. The two parallel factors are defined in part by tests employing syllogistic items. For ESI, Symbol Reasoning employs symbolic-syllogistic items and for EMI, Logical Reasoning employs the more usual verbal-syllogistic items. Both have some tendency toward factorial complexity. But whereas Logical Reasoning has a significant loading on EMR, for which it was formerly believed to be one of the best measures, Symbol Reasoning does not have a parallel significant loading on ESR. This may be because the given premises in the latter do not contain simple relations.

Other parallels between tests could be pointed out, but this could become tedious. One more will be mentioned, however, a case in which the appearance of test similarity is not so great, but the psychological similarity is good. This case is for the test Letter "U" for factor ESU and the test Double Descriptions for factor EMU. Both involve specifying properties of the kind of unit to be used as the "model"; i.e., its class is specified. In Letter "U", E is to decide which words satisfy the specification "words containing the letter 'U'." In Double Descriptions, E is to decide which object most nearly satisfies two properties, such as

both white and edible. Deciding whether or how well given objects fit the specifications of the class is the crucial part of either task.

But one can also cite instances in which seemingly parallel tests do not both measure two parallel evaluation factors, even factors ESU and EMU. The test Symbol Identities is the leading measure of ESU. It calls for decisions as to whether two letter or number sets are identical in every detail. Two tests intended for EMU emphasized, not identities of word meanings but very near similarities of word meanings (Nihira, et al., 1964). One was a matter of deciding which synonym comes nearest in meaning to a given word, when words are given out of context, and the other put the stem word for the item within the context of a sentence. Both tests failed to have significant relations with EMU and had significant loadings on CMU instead. Such discrepancies call for further investigation. It would appear that some of the same testing principles cannot be effectively applied as we change kind of content or kind of product within an area of evaluation.

New Light on the Properties of Evaluation

What can it be said that the study has done for further elucidation of the concept of the operation of evaluation? What aspects of the concept possibly need changing as a result of the new information about factors and their tests, and what new features come into the picture? We shall consider the answers to these general questions more specifically in terms of (1) the kinds of judgments that belong in the picture of evaluation as defined by the tests; (2) what kinds of criteria for judgment are pertinent to measurement of the evaluation abilities; (3) whether remembered information can be utilized in evaluation; and (4) whether there are any new restrictions to be placed on the concept. The answers to these issues rest upon the kinds of tests that serve to measure their evaluation factors well and those that do not, when all have been hypothesized to do so.

Kinds of Judgment

Much was said from one place to another in this report about the two classes of tests: sensitivity vs. estimation. More fully spelled out, these terms mean sensitivity to error or discrepancy on the one hand and, on the other, judgment of relative nearness of a number of items of information (for any kind of product) to a kind of model item of information on the same continuum. In more operational terms, the contrast may be stated in terms of "absolute" vs. "relative" kinds of judgments, as in psychophysics.

It should be abundantly clear, from the way in which both sensitivity and estimation types of tests are commonly related significantly to the same factors, that both kinds of judgments apply. If we compare evaluation tests from the two categories that clearly involve absolute vs. relative judgments, however, we find that the 11 requiring absolute judgments tend to have slightly stronger loadings on the factors for which they were intended than the 12 requiring relative judgments, the average difference being about .07. Four of the 12 tests involving relative judgments had higher loadings on factors for which they were not intended, only one of the four having significant loading on its intended factor, the other such loadings being .01, .09, and .20. This compares with two of the 11 tests involving absolute judgments that had higher loadings on factors for which they were not intended, with loadings of .35 and .29 on factors for which they were intended.

Although these differences are in all probability not significant, they suggest a trend. On the whole, tests requiring absolute judgments are operationally simpler, which should be more favorable for univocality. The best examples of such simple tests are Perceptual Speed and Identical Forms, for factor EFU, and Symbol Identities and Letter "U", for factor ESU.

Kinds of Criteria

The most common criteria employed in evaluation tests have been identity vs. non-identity and consistency vs. inconsistency. Other kinds of criteria have been mentioned in connection with various tests and for different factors. Some examples are: fitness for class membership (Letter "U"); relative familiarity (Familiar Letter Combinations); relative similarity (Series Relations); conformity to principles (Correct Number Series, Correct Letter Orders); relative probability (Decoding); and solvability of problems (Letter Problems). Such a variety of criteria can possibly be brought under more abstract and more general criterion categories, since such terms as "identity," "similarity," and "conformity" suggest that the criteria tend to be logical in nature and that they represent continua of one kind or another.

Perceived vs. Remembered Models

There were two tests designed for factor ESU that failed to function as expected — Correct Spelling and Familiar Letter Combinations. A reason offered for the failures was to the effect that in both instances the examinee has to draw upon his memory storage for something with which to compare the given (perceived) item of information. This suggested the generalization that in evaluation activities the comparisons must be between items of information both present to cognition, if not to perception, at the moment. No other evaluation tests offered the possibility of finding clear evidence on this point, thus leaving the question open. From a priori thinking, it would seem that in everyday life we are perpetually comparing presently cognized information with remembered information. There would seem to be good possibilities for testing the hypothesis experimentally.

The Scope of Evaluation

The scope of processes under the heading of evaluation is indicated somewhat by the variety of criteria that may be involved. The discussion of this topic above revealed something of the apparent variety of criteria that is represented in the experimental tests. But it was suggested that such criteria may be limited to the general logical category, which would rule out of consideration criteria involving ethical and esthetic values. There is no doubt that such values exist and such areas of judgment call for evaluative operations. At the present, they do not seem to fit into the structure-of-intellect model. Perhaps they call for two complete additional sets of evaluative abilities or processes, whether parallel to the present theoretical set or not.

As to the definition of evaluation itself, the kind of restriction just discussed suggests that it is going too far to say that evaluation is a matter of reaching decisions regarding goal satisfaction. Kinds of goals are much too numerous, and satisfaction in terms of logical criteria cannot cover all cases. In defining the restricted kinds of evaluation represented in the structure of intellect, it would seem desirable to eliminate reference to "goal satisfaction."

As a general impression, from consideration of the experimental tests and their factors in this study, the importance of an act of comparison seems to stand out. This observation was also made by Hertzka (1953) when he said, "The core of this definition is the concept of comparison." The following current definition of evaluation can be suggested: Evaluation is a process of comparing a product of information with known information according to logical criteria, reaching a decision concerning criterion satisfaction.

Tests Recommended for Symbolic-Evaluation Factors

The following tests demonstrate a sufficient degree of univocality and saturation of variance to merit being recommended as tests of their respective factors in future research.

ESU	36.	Symbol Identities
	19.	Letter "U"
ESC	32.	Sign Changes II
	3.	Best Number Class
ESR	37.	Symbol Manipulation
	27.	Related Words I
ESS	30.	Series Relations
	41.	Way-Out Numbers
EST	16.	Jumbled Words
ESI	1.	Abbreviations

The authors do not mean to imply that these tests are to be accepted unquestionedly as the ultimate measures of their respective factors, but rather that these tests appear to be the best measures available at present. All the tests await further refinement and analysis before it can be said with certainty that any one of them is the best possible measure of its factor.

SUMMARY

The study was designed to test the implications of, and extend the empirical foundations underlying the structure-of-intellect model. The six hypothesized abilities of symbolic evaluation were selected for investigation. Abilities in this category appear to be important for mathematical thinking.

Five major objectives of the study were: (1) Demonstration of the six symbolic-evaluation factors hypothesized by the structure-of-intellect model. (2) Investigation of what mental processes are evaluative. (3) Investigation of what materials are processed symbolically. (4) Clarification of selected known factors of the model through the use of shortened and refined tests. (5) Evaluation of the theoretical model as a basis for predicting still-unknown factors and for developing tests of these factors.

The six experimental factors were: evaluation of symbolic units (ESU), evaluation of symbolic classes (ESC), evaluation of symbolic relations (ESR), evaluation of symbolic systems (ESS), evaluation of symbolic transformations (EST), and evaluation of symbolic implications (ESI).

In order to demonstrate the distinctness of the hypothesized factors from those already known, 12 reference factors, previously confirmed, were analyzed as experimental controls: cognition of symbolic units (CSU), cognition of symbolic classes (CSC), cognition of symbolic relations (CSR), cognition of symbolic systems (CSS), cognition of symbolic implications (CSI), cognition of semantic units (CMU), memory for symbolic implications (MSI), divergent production of symbolic classes (DSC), convergent production of symbolic systems (NSS), convergent production of symbolic transformations (NST), convergent production of symbolic implications (NSI), and evaluation of figural units (EFU).

A thirteenth factor, not formally hypothesized by the structure-of-intellect model, was included in order to determine its possible relationships to performance on intellectual tests: motivation to succeed (MOT).

In order to accomplish the experimental objectives, 53 test measures were constructed, adapted, or selected to compose an eight-hour battery. Twenty-five tests were designed to measure the six experimental factors, the remaining tests were adapted marker tests for reference factors, ratings designed to reflect motivational level, and published tests from school records.

The battery was administered to 225 high-school seniors in a Southern California community. The scores were intercorrelated and factor analyzed. The twenty principal-axis factors were rotated both graphically and analytically, observing the criteria of simple structure, positive manifold, and psychological meaningfulness.

All six hypothesized factors emerged, defined largely by the tests designed to measure them. Twelve reference factors were isolated, indicating that the shortened, refined tests remained relatively factorially invariant over several analyses and major revisions. A general motivation factor failed to emerge. Instead, a rating doublet and prediction-achievement-discrepancy singlet were found.

Among the six experimental factors, ESU was clearly defined as a factor that could be described as clerical speed and accuracy, the ability to judge rapidly symbolic material in terms of identity or error. The ESC factor was least clearcut, but it involved sensitivity to class properties. ESR was clearly isolated as the ability to make choices among symbolic relationships on the bases of identity and consistency. The ESS factor appeared to involve the estimation of similarity among series. The clear isolation of EST defined it as the ability of sensitivity to the identical nature of rearrangements and substitutions of letters within words. The ability to judge the consistency or probability of implications from symbolic material was represented by the ESI factor.

From the tendency of the hypothesized factors to be defined by tests involving sensitivity or estimation (absolute or relative judgments), it is concluded that the two are to be considered functionally equivalent in evaluative operations.

This study further confirmed the established finding that numbers, letters, and words can be processed as symbolic information.

Relatively clear separation among abilities according to operation, content, and product was interpreted as substantially contributing to the value of the unified theoretical model of the structure of intellect as a hypothetico-deductive theory for generating predictions concerning individual differences in intellectual functioning.

REFERENCES

Bechtoldt, H. P. Factorial investigation of the perceptual speed factor. Amer. Psychologist, 1947, 2, 304-305.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. Differential aptitude tests. (3rd ed.) New York: The Psychological Corporation, 1959.

Berger, R. M., Guilford, J. P., & Christensen, P. R. A factor-analytic study of planning abilities. Psychol. Monogr., 1957, 71, (Whole No. 435).

Boring, E. G, Intelligence as the tests test it. The New Republic, June 6, 1923, 35-37.

Canisia, Sister M. Mathematical ability as related to reasoning and use of symbols. Educ. psychol. Measmt, 1962, 22, 105-127.

Cattell, R. B. A universal index for psychological factors. Adv. publ. No. 3, laboratory of personality assessment as group behavior. Urbana: University of Illinois, 1953.

Christensen, P. R., & Guilford, J. P. An experimental study of verbal fluency factors. Brit. J. stat. Psychol., 1963, 26, 1-26.

Cliff, N. Orthogonal rotations to congruence. Amer. Psychologist, 1964, 19, 582. (Abstract)

Coombs, C. H. A factorial study of number ability. Psychometrika, 1941, 6, 161-189.

Davis, P. C. A factor analysis of the Wechsler-Bellevue scale. Educ. psychol. Measmt, 1956, 16, 127-146.

deMille, R. Intellect after lobotomy in schizophrenia. Psychol. Monogr., 1962, 76, (Whole No. 535).

Educational Testing Service. Cooperative school and college ability test (Form 2A). Princeton: Author, 1958.

Educational Testing Service. Preliminary scholastic aptitude test, grades 11-12. Princeton: Author, 1959.

French, J. W. The description of aptitude and achievement tests in terms of rotated factors. Psychometr. Monogr., 1951, No. 5.

French, J. W., Ekstrom, Ruth B., & Price, L. A. Manual for kit of reference tests for cognitive factors. Princeton: ETS, 1963.

Gershon, A., Guilford, J. P., & Merrifield, P. R. Figural and symbolic divergent-thinking abilities in adolescent and adult populations. Rep. psychol. Lab., No. 29. Los Angeles: Univer. Southern Calif., 1963.

Green, R. F., Guilford, J. P., & Christensen, P. R. A factor-analytic study of reasoning abilities. Psychometrika, 1953, 18, 135-160.

Guilford, J. P. Basic conceptual problems in the psychology of thinking. Ann. N. Y. Acad. of Sci., 1960, 19, 6-21.

Guilford, J. P. Basic problems in teaching for creativity. Paper read at the Conference on Creativity and Teaching Media, La Jolla, Calif., September, 1964.

Guilford, J. P., Christensen, P. R., Kettner, N. W., Green, R. F., & Hertzka, A. F. A factor-analytic study of Navy reasoning tests with the Air Force Aircrew Classification Battery. Educ. psychol. Measmt, 1954, 14, 301-325.

Guilford, J. P., & Hoepfner, R. Current summary of structure-of-intellect factors and suggested tests. Rep. psychol. Lab., No. 30. Los Angeles: Univer. Southern Calif., 1963.

Guilford, J. P., & Lacey, J. I. (Eds.) Printed classification tests. A. A. F. Aviation Psychology Research Program Reports, No. 5. Washington, D. C.: GPO, 1947.

Guilford, J. P., & Merrifield, P. R. The structure of intellect model: its uses and implications. Rep. psychol. Lab., No. 24. Los Angeles: Univer. Southern Calif., 1960.

Guilford, J. P., Merrifield, P. R., Christensen, P. R., & Frick, J. W. Some new symbolic factors of cognition and convergent production. Educ. psychol. Measmt, 1961, 21, 515-541.

Guilford, J. P., & Zimmerman, W. S. The Guilford-Zimmerman aptitude survey. Beverly Hills: Sheridan Supply Co., 1956.

Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.

Hertzka, A. F. A factorial analysis in the domain of evaluation. Unpublished doctoral dissertation, University of Southern California, 1953.

Hertzka, A. F., Guilford, J. P., Christensen, P. R., & Berger, R. M. A factor-analytic study of evaluative abilities. Educ. psychol. Measmt, 1954, 14, 581-597.

Johnson, D. M. The psychology of thought and judgment. New York: Harper, 1955.

Kettner, N. W., Guilford, J. P., & Christensen, P. R. A factor-analytic study across the domains of reasoning, creativity, and evaluation. Psychol. Monogr., 1959, 73, (Whole No. 479).

Lindquist, E. F. Iowa tests of educational development. Chicago: Science Research Associates, 1959.

Nihira, K., Guilford, J. P., Hoepfner, R., & Merrifield, P. R. A factor analysis of the semantic-evaluation abilities. Rep. psychol. Lab., No. 32. Los Angeles: Univer. Southern Calif., 1964.

Pemberton, C. L. The closure factors related to other cognitive processes. Psychometrika, 1952, 17, 267-288.

Petersen, H., Guilford, J. P., Hoepfner, R., & Merrifield, P. R. Determination of structure-of-intellect abilities involved in ninth-grade algebra and general mathematics. Rep. psychol. Lab., No. 31. Los Angeles: Univer. Southern Calif., 1963.

Restle, F. Psychology of judgment and choice: a theoretical essay. New York: Wiley, 1961.

Thurstone, L. L. Primary mental abilities. Psychometr. Monogr., 1938, No. 1. (a)

Thurstone, L. L. The perceptual factor. Psychometrika, 1938, 3, 1-17. (b)

Thurstone, L. L. Experimental study of simple structure. Psychometrika, 1940, 5, 153-168.

Underwood, B. J., & Schulz, R. W. Meaningfulness and verbal learning. Philadelphia: Lippincott, 1960.

Vinacke, W. E. The psychology of thinking. New York: McGraw-Hill, 1952.

Vinacke, W. E. Relations between motivational conditions and thinking. Ann. N. Y. Acad. of Sci., 1960, 19, 76-93.

Wilson, R. C., Guilford, J. P., & Christensen, P. R. A factor-analytic study of creative-thinking abilities. Psychometrika, 1954, 19, 297-311.

APPENDIX

DESCRIPTION OF TESTS

1. Abbreviations - ESI01B. Choose one of three alternative words that a given abbreviation best implies.

Sample Item: crnt
A. crescent
B. coronation
C. current

Answer: C

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

2. Best Letter Set - ESS03A. Choose one of three alternative letter sets that is most like the given set.

Sample Item: EKN
A. JFI
B. PAQ
C. YBT

Answer: C (begins with a vowel)

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 10 minutes.

3. Best Number Class - ESC01A. Judge into which class a given number fits so as to receive the most possible points. The classes and their points are: EVEN MULTIPLES - 1 point; ODD MULTIPLES - 2 points; SQUARES - 3 points; PRIMES - 4 points.

Sample Item: 100

Answer: Class of SQUARES

Score: Number right minus one-third number wrong. Parts: 2; items per part: 15; working time: 6 minutes.

4. Best Number Pairs - ESC02A. Choose one of three number pairs that makes the best class. In order, from best to worst, the classes are: perfect squares, multiples, odd-or-evens, and no property in common.

Sample Item:
A. 6 - 4
B. 4 - 9
C. 9 - 6

Answer: B (perfect squares)

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

5. Camouflaged Words - NST01A. Find within a meaningful sentence a group of consecutive letters that, in the given order, spell the name of a sport or game.

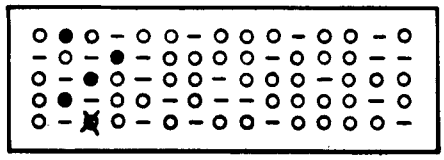
Sample Item: COWARDICE IS NOT A SOLDIERLY ATTRIBUTE

Answer: DICE

Score: Number of correct responses, one camouflaged word per item. Parts: 1; items: 15; working time: 6 minutes.

6. Circle Reasoning[a] - CSS01D. Discover the principle by which one circle is blackened in each of four rows of circles and dashes. Apply the rule to the fifth row.

Sample Item:



Answer: The second circle from the left is blackened.
Score: Number of correct responses. Parts: 1; items: 14; working time: 8 minutes.

7. Correct Letter Orders - ESS04A. Judge whether or not the order of given letters follows a given rule.

Sample Items: Rule - Alternate letters in the alphabet (skipping one).

I. M O Q S U W
II. P R S U W Y

Answer: I follows the rule; II does not follow the rule.

Score: Number right minus number wrong. Parts: 2; items per part: 20; working time: 6 minutes.

8. Correct Number Series - ESS05A. Judge whether or not the order of given numbers follows a given rule.

Sample Items: Rule - Alternately add 1, multiply by 3.

I. 2 3 9 10 30 31
II. 4 5 15 16 49 50

Answer: I follows the rule; II does not follow the rule.

Score: Number right minus number wrong. Parts: 2; items per part: 20; working time: 6 minutes.

9. Correct Spelling - ESU04A. Judge whether or not given words are spelled correctly.

Sample Items:
I. experience
II. seperate

Answer: I correct; II incorrect.

Score: Number right minus number wrong. Parts: 2; items per part: 60; working time: 6 minutes.

[a] Adapted with permission from a test by Robert Blakey.

10. Decoding - EST01A. Choose the word that would be easier to decode if coded. The code is based upon probabilities of letter occurrences and is a highly ambiguous code.

Code:
All double letters (oo, gg, etc.) are 1.
All pairs of vowels (ea, ou, etc.) are 2.
All pairs of consonants (bl, sh, etc.) are 3.
All vowels (a, e, i, o, u, y) are 4.
All consonants (b, g, p, etc.) are 5.

Sample Item:
A. little
B. blood
C. both words

Answer: B

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

11. Derivations - ESU08A. Rapidly judge whether words can be derived from a given word by using some of its letters.

Sample Items: Given: PROCRASTINATE

I. trap
II. crust
III. percent

Answer: I can be derived; II and III cannot.

Score: Number right minus number wrong. Parts: 3; items per part: 50; working time: 9 minutes.

12. Disemvowelled Words[b] - CSU04B. Recognize familiar words with dashes in place of vowels; then complete the words by writing the vowels.

Sample Item: m__t__l__t__ Answer: mutilate

Score: Number of correctly completed words. Parts: 1; items: 25; working time: 5 minutes.

13. Familiar Letter Combinations - ESU05A. Choose one of two given letter combinations that is the more familiar syllable.

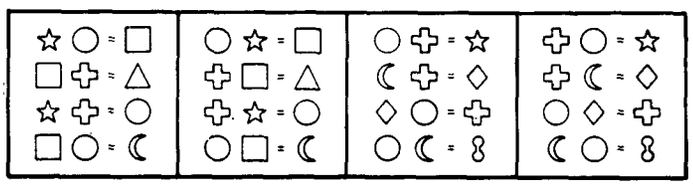
Sample Item:
A. l o y
B. n i s

Answer: A

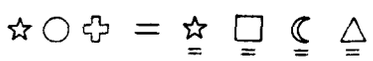
Score: Number right minus number wrong. Parts: 2; items per part: 20; working time: 6 minutes.

14. Form Reasoning[a] - NSI02C. From the table, find the form that is equal to the three given forms.

Table:



Sample Item:

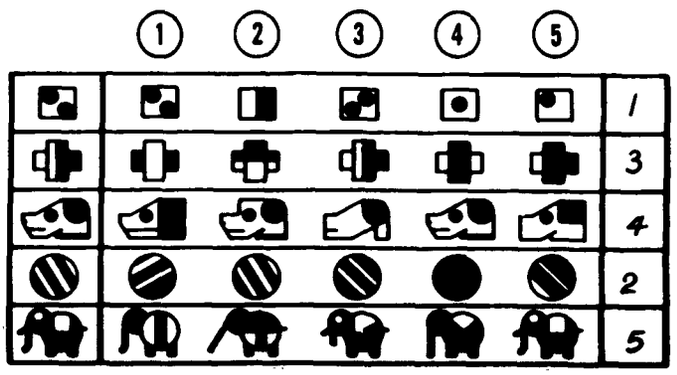


Answer:

Score: Number right minus one-third number wrong. Parts: 1; items: 20; working time: 4 minutes.

15. Identical Forms[c] - EFU02A. Find one of five figures that is exactly the same as the given figure.

Sample Items:



Score: Number right minus one-fourth number wrong. Parts: 1; items: 60; working time: 3 minutes.

16. Jumbled Words - EST03A. Judge whether or not words could be made by mixing the letters of a given word.

Sample Items: Given word: start

I. stare
II. stars
III. tarts

Answers: I and II can't be made; III can be made.

Score: Number right minus number wrong. Parts: 2; items per part: 25; working time: 4 minutes.

[b] Adapted with permission from a test by James J. McGrath.
[c] Adapted with permission from a test by L. L. Thurstone.

17. Letter Problems - ESI02A. From a table of letter-problem solutions, judge whether larger given problems can be solved directly, solved with transpositions, or are unsolvable.

Table:

T Z = U Z Y = V X X = Z	Y Y = W T V = Z Z U = W	Y X = T V W = Y U U = T

Sample Items: I. T Z U
II. U V Z
III. X Z X

Answers: I can be solved directly; II can be solved only with transpositions; and III is unsolvable.

Score: Number right minus one-half number wrong. Parts: 3; items per part: 10; working time: 12 minutes.

18. Letter Triangle - CSS02B. Find the pattern of the letters arranged systematically within a triangle.

Sample Item:

a
b c
d e f
? — — —

A. h
B. j
C. g
D. f
E. a

Answer: C (g)

Score: Number right minus one-fourth number wrong. Parts: 1; items: 12; working time: 9 minutes.

19. Letter "U" - ESU06A. Check all the words in lists that contain the letter "u".

Sample Items:
( ) sense
( ) short
(X) juice
( ) special

Score: Number correctly placed X's minus number incorrectly placed X's. Parts: 2; items per part: 200; working time: 2 minutes.

20. Marking Speed Test. Make as many X's as you can in the rows of squares provided.

Sample Item:

Score: Number of complete X's made within squares. Parts: 1; items: 180 squares; working time: 1 minute.

21. Number Classification - CSC03B. Select one of five alternative numbers to fit into each of four classes of three given numbers each.

Sample Items:
I. 44 55 33
II. 10 45 15

A. 421
B. 53
C. 219
D. 22
E. 25

Answers: I, D; II, E.

Score: Number right minus one-fourth number wrong. Parts: 1; items: 16; working time: 5 minutes.

22. Number Grouping - DSC01B. Group given numbers into several different classes based upon properties they have in common, e.g., multiples of three.

Sample Item: Given: 2 3 4 6 17 23 36

Possible groups: a. 2, 4, 6, 36 (even numbers)
b. 3, 17, 23 (odd numbers)

Score: Number of groups listed that exhaust their members. Parts: 2; items per part: 3; working time: 6 minutes.

23. Number-Group Naming - CSC05A. State how the numbers in each set of three are alike.

Sample Item: 35 110 75 Answer: Divisible by 5 (or multiples of 5)

Score: Number of correctly named groups. Parts: 1; items: 12; working time: 3 minutes.

24. Numerical Operations[d] - MSI01B. Part III of the Guilford-Zimmerman Aptitude Survey (adapted).

Rapidly add, subtract, or multiply simple numerical problems and select one of six alternatives as the answer.

Score: Number right minus one-fifth number wrong. Parts: 2; items per part: 36; working time: 4 minutes.

25. Operations Sequence - NSS01B. Produce the correct order of three specified numerical operations in order to get from one given number to another.

Sample Item: Start with 6 obtain 18

a. + 3
b. ÷ 2
c. x 3

Answer: b a c

Score: Number of correct orders minus one-fourth number of incorrect orders. Parts: 2; items per part: 12; working time: 10 minutes.

26. Perceptual Speed[d] - EFU01A. Part IV of the Guilford-Zimmerman Aptitude Survey (adapted).

Rapidly match each of five objects to one of four given objects.

Score: Number right minus one-fourth number wrong. Parts: 1; items: 72; working time: 5 minutes.

[d] Adapted with permission from Sheridan Supply Co., Beverly Hills, Calif.

27. Related Words I - ESR03A. Choose the alternative word pair with a relation most like that of the given pair.

Sample Item: GRAND - RAN

A. country - cot
B. respite - sit
C. loving - log

Answer: B

Score: Number right minus one-half number wrong. Parts: 3; items per part: 10; working time: 12 minutes.

28. S Test - ESI04A. Discover problems in items composed of numbers, letters, and words and solve the problems.

Score: Number of problems solved which are symbolic in nature. Parts: 1; items: 20; working time: 5 minutes.

29. Seeing Trends II - CSR01B. Describe a trend based upon relations of letters in a group of words.

Sample Item: rated crate morning dearth separate

Answer: The "r" moves one letter to the right in each word.

Score: Number of correct descriptions. Parts: 1; items: 15; working time: 10 minutes.

30. Series Relations - ESS06A. Choose one of three arithmetic operations that best relates each number in a given series to the previous number.

Sample Item: 17, 9, 2 (given series)

A. - 8
B. ÷ 2
C. - 7

Answer: A (none is correct, but A is best)

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

31. Sign Changes - NSI01A. Solve simple arithmetic problems in which the operation sign is changed according to a set of rules.

Sample Item: Rules: wherever you see: - replace it by x
+ replace it by -

3 - 6 = 18
6 + 2 = 4
4 - 3 = 12

Score: Number right solutions. Parts: 1; items: 32; working time: 1 minute.

32. Sign Changes II - ESR01C. Choose the sign changes that make the expressions into equations.

Sample Item: 3 + 1 = 6 x 2

A. Instead of + you -
B. Instead of + you x
C. Instead of x you -
D. Both A and C

Answer: C

Score: Number right minus one-third number wrong. Parts: 2; items per part: 10; working time: 8 minutes.

33. Similar Pairs - ESR04A. Judge whether or not the relation in the second pair of words is the same as the relation in the first pair.

Sample Items:

I.	kire - lire	fora - gora
II.	brake - rake	freed - reed
III.	moan - noam	toes - seot

Answers: I and II have same relation; III does not.

Score: Number right minus number wrong. Parts: 2; items per part: 15; working time: 6 minutes.

34. Sound Grouping - ESC04A. Find the word that does not belong in a group of four words.

Sample Item:
A. comb
B. foam
C. home
D. come

Answer: D

Score: Number right minus one-third number wrong. Parts: 2; items per part: 15; working time: 6 minutes.

35. Symbol Grouping - CSI01B. Rearrange scrambled symbols in a specified systematic order as efficiently as possible.

Sample Item: Given: x-ox-x By moving one or more adjacent symbols at a time, arrive at the order xxx--o.

Answer: (the minimum number of moves is two)

Score: Two points for the most efficient rearrangement; one point for a less efficient rearrangement; and no points for an inefficient or incorrect rearrangement. Parts: 1; items: 12; working time: 7 minutes.

36. Symbol Identities - ESU07A. Judge whether both members of pairs of words and numbers are the same or are different.

Sample Items:

2 1 6 3	S	2 1 6 3
Hahn, Lorena	D	Hahn, Lorina
Bob Ulm	D	Rob Ulm

Score: Number right. Parts: 2; items per part: 50; working time: 4 minutes.

37. Symbol Manipulation - ESR02C. Judge whether symbolic conclusions are true or false based upon given premises.

Sample Items: If X is smaller than Y (X s Y), then:

I. X e Y (X equals Y)
II. X nl Y (X is not larger than Y)
III. X l Y (X is larger than Y)

Answers: I and III are false; II is true.

Score: Number right minus number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

38. Symbol Reasoning - ESI03A. Decide whether symbolically stated conclusions are true, false, or uncertain, based upon a given symbolic relationship.

Sample Items: Given: $2x < 3y < 2z$

		Answers:	
I.	$2x = 2z$	I.	false
II.	$y < z$	II.	true
III.	$x = y$	III.	uncertain

Score: Number right minus one-half number wrong. Parts: 2; items per part: 24; working time: 12 minutes.

39. Typing Errors - EST02A. Choose one of three correctly typed words that the given incorrectly typed word would most likely be. Typing rules and keyboard are printed on the test.

Sample Item: F H E E

A. thee
B. tree
C. free

Answer: A

Score: Number right minus one-half number wrong. Parts: 2; items per part: 15; working time: 10 minutes.

40. Varied Symbols - DSC03B. Find the different common properties that sets of letter combinations may have in common.

Sample Item: The set | EPZT | APCTO | UMDT | is like which of these groups:

1. ACBE
2. ROS
3. COM
4. GAIH
5. ZMOD

Possible answers: 1 (start with vowels); 5 (have three consonants)

Score: Number of indicated groups having common properties. Parts: 2; items per part: 5; working time: 8 minutes.

41. Way-Out Numbers - ESS07A. Choose the one alternative number from a list of four ordered numbers that is farthest away from the other three.

Sample Item:

A. 31
36
45
B. 47

Answer: A

Score: Number right minus number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

42. Word Changes - NSS02C. Arrange a list of words, each containing the same number of letters, so that the first word is changed into the last word with only one letter change at each step.

Sample Item:

BELL
____
____
MAIN

1. BAIL
2. BALL
3. MAIL

Answer: 2, 1, 3

Score: One point each for correct order of first five items; two points each for correct order of last five items. Parts: 1; items: 10; working time: 6 minutes.

43. Word Choice - ESC03A. Choose one of three alternative words that best fits the given class of words. The word classes are based upon common properties of letters.

Sample Item:

school
fleet
doomsday

A. delete
B. relate
C. expect

Answer: A (none are correct, but A is best)

Score: Number correct minus one-half number wrong. Parts: 2; items per part: 15; working time: 10 minutes.

44. Word Combinations - CSU06A. Produce a new word from the ending of one word and the beginning of another.

Sample Items: Given:

1. bridge
2. beam
3. open

A. duress
B. zero
C. pledge
D. need
E. none of these

Answers:

1. D (gene)
2. C (ample)
3. A (endure)

Score: Number right minus one-fourth number wrong. Parts: 2; items per part: 15; working time: 8 minutes.

45. Word Patterns - CSI03C. Arrange a list of short words efficiently in a crossword-puzzle design.

Sample Item: Given words: bats, easy, hot, tea, the

Answer: most efficient arrangement:

b
t e a
h o t
e a s y

Score: Reciprocal of the total number of inefficient and incorrect arrangements of letters and words multiplied by 100. Parts: 1; items: 6; working time: 12 minutes.

46. Word Relations - CSR02B. Recognize the same relation between words in each of two pairs, then complete a third pair from five alternative words using the same relation.

Sample Item:
on - no
top - pot
part - ?

A. art
B. pat
C. rapt
D. tar
E. trap

Answer: E

Score: Number right minus one-fourth number wrong. Parts: 2; items per part: 10; working time: 6 minutes.

47. Word Transformation - NST02B. Separate letters of words in a phrase with vertical lines to make a different set of words.

Sample Item (answer marked): RING S|O|FT HE

Score: Number of vertical lines regrouping all the letters in the given phrase. Parts: 1; items: 15; working time: 4 minutes.

48. Rating of Test-liking. Rate how much you liked the last test on a five-point scale. The five categories of rating are: very much; pretty much; a little; not very much; not at all.

Score: Sum of standardized ratings for all 47 tests.

49. Rating of Effort Spent on Test Booklet. Rate how hard you worked on this booklet on a five-point scale. The five categories of rating are: as hard as I could; very hard; fairly hard; not very hard; not hard at all.

Score: Sum of standardized ratings for all 8 booklets.

50. Sex Membership. Score: Boys were given a score of 1; girls a score of 0.

51. Achievement-Prediction Discrepancy Score. Score: Four-year high-school grade-point average minus the least-squares prediction of that average by the eight subtests of the Differential Aptitudes Test.

52. Iowa Tests of Educational Development - Test 8, General Vocabulary - CMU. Recognize the meanings of words commonly used in communication. This test is similar to standard verbal comprehension (CMU) marker tests.

53. Preliminary Scholastic Aptitude Test - Verbal - CMU. PSAT is an abbreviated form of the SAT. The verbal score is the sum of scores on four tests: Opposites, Sentence Completion, Analogies, and Reading Comprehension. The dominant saturation is hypothesized to be CMU with some CMR variance contributed by the Analogies test.

54. Cooperative School and College Ability Test, Form 2A - Verbal - CMU. Verbal score is composed of scores on two tests, Sentence Understanding and Word Meanings. Tests similar to each have previously loaded on the verbal comprehension (CMU) factor.

55. Mathematics Experience. A weighted total of the number of high-school mathematics courses completed and currently being taken. One point was given for each semester of the following courses: Consecutive Mathematics (grades 5-8), Remedial Arithmetic, Preparatory Algebra Arithmetic, Algebra I, and Geometry I, and two points for each semester of Modern Algebra, Introductory Algebra-Trigonometry, Modern Intermediate Mathematics, Modern Geometry, and College Mathematics.

56. Typing Experience. Total number of semesters of high-school typing courses completed and currently being taken. One point was given for each semester of the following courses: Junior-High Typing, Typing I, and Typing II.

57. Shorthand Experience. Total number of semesters of high-school shorthand courses completed and currently being taken. One point was given for each semester of Shorthand I and Shorthand II.

58. Test Administrator. Score: 150 examinees tested by one examiner were given scores of 0; 75 tested by another examiner were given scores of 1.

Unclassified
Security Classification

DOCUMENT CONTROL DATA · R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Aptitudes Research Project Department of Psychology University of Southern California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP ---

3. REPORT TITLE

A FACTOR ANALYSIS OF THE SYMBOLIC - EVALUATION ABILITIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (Last name, first name, initial)
Hoepfner, Ralph
Guilford, J. Paul
Merrifield, Philip R.

6. REPORT DATE December, 1964	7a. TOTAL NO. OF PAGES 39	7b. NO. OF REFS 42
8a. CONTRACT OR GRANT NO. Nonr-228(20) b. PROJECT NO. NR 150-044 c. --- d. ---	9a. ORIGINATOR'S REPORT NUMBER(S) Reports from the Psychological Laboratory University of Southern California Number 33. 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ---	

10. AVAILABILITY/LIMITATION NOTICES
U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through originating agency.

11. SUPPLEMENTARY NOTES ---	12. SPONSORING MILITARY ACTIVITY Personnel and Training Branch Office of Naval Research Department of the Navy

13. ABSTRACT

This study approaches the problem of describing symbolic judgmental processes from the standpoint of individual differences in terms of basic traits. Based upon Guilford's structure-of-intellect model, six factors of symbolic evaluation were hypothesized to exist as distinct from each other and also from factors represented in certain other domains of the model. (U)

Twenty-five experimental tests were developed to serve as measures of the six hypothesized factors. In order to obtain factors clearly generalizable to those generated from the model, at least three tests were developed for each hypothesized factor. Furthermore, the tests designed were not specific to a certain kind of symbol but utilized several kinds, including numbers, words, letters, and syllables. (U)

Measures of reference factors from the structure of intellect were also included in order to demonstrate clearly the uniqueness of the hypothesized factors. Twenty-five tests designed to measure 12 reference factors were analyzed as part of the total battery. In addition, three variables designed to measure the examinees' levels of motivation to succeed on the aptitude tests were employed in expectation of isolating a motivation factor and possibly determining the influence of motivation upon the test scores. (U)

All the tests were administered to 225 high-school seniors in a Los Angeles County school. The correlation matrix was factor analyzed and axes were analytically rotated to a least-squares fit to a simple-structure target matrix. Twenty factors were isolated and named, including the six hypothesized symbolic-evaluation factors, the twelve reference factors, and two different factors defined by the motivation measures. (U)

The general conclusion of this research is that the structure-of-intellect model has continued to lead fruitfully to undiscovered, differentiable intellectual abilities. More specifically, five new factors were added to the list of recognized intellectual abilities and others have been verified. (U)

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Intelligence Tests Factor Analysis Evaluation						

DD FORM 1 JAN 64 1473 0101-807-6800

Unclassified
Security Classification